



2-24-00

412 Rec'd PCT/PTO 23 FEB 2000

146.1335

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Application of: :
JESUS CORTES et al :
PCT No.: PCT/FR98/01593 :
Filed: January 25, 2000 :
For: BIOSYNTHESIS...AND THEIR USE :

PCT Date: July 2, 2000

600 Third Avenue
New York N.Y. 10017
February 23, 2000

COMPLETION OF APPLICATION

Asst. Commissioner for Patents
Washington, D.C. 20231

Sir:

In order to complete the filing of the above-application, Applicants are submitting herewith the executed declaration to replace the unexecuted declaration filed with the application. Also enclosed is a check for \$130.00 for the completion fee. Since the filing of the application is now complete, it is requested that it be forwarded to the group for examination in due course.

Respectfully submitted,
Bierman, Muserlian and Lucas

By:

Charles A. Muserlian
Charles A. Muserlian #19,683
Attorney for Applicants
Tel.# (212) 661-8000

03/01/2000, PVOLPE 00000019 09463705

01 FC:154

130.00 OP

CAM:ds
Enclosures

Express Mail[®] money order
482/5992
Feb. 23, 2000
I hereby certify that this paper or fee is
being deposited with the United States Postal
Service "Express Mail Post Office to Addressee"
service under 27 CFR 1.10 on the date indicated
above and is addressed to the Commissioner
of Patents and Trademarks, Washington, D.C.
20540.

Feb. 23, 2000

D. Falcone
Signature of person making paper of fee

FORM PTO-1199 (REV. 5-93)		U.S. DEPARTMENT OF COMMERCE PATENT AND TRADEMARK OFFICE	ATTORNEYS DOCKET NUMBER 146.1335
TRANSMITTAL LETTER TO THE UNITED STATES DESIGNATED/ELECTED OFFICE (DO/EO/US) CONCERNING A FILING UNDER 35 U.S.C. 371			U.S. APPLICATION NO. 09/463705
INTERNATIONAL APPLICATION NO. PCT/FR98/01593	INTERNATIONAL FILING DATE July 21, 1998	PRIORITY DATE CLAIMED July 25, 1997	
TITLE OF INVENTION: BIOSYNTHESIS GENES AND TRANSFER OF 6-DESOXY-HEXOSES IN SACCHAROPOLYSPORA ERYTHRAEA AND IN STREPTOMYCES ANTIBIOTICUS AND THEIR USE			
APPLICANT(S) FOR DO/EO/US FROMENTIN et al			
Applicant herewith submits to the United States Designated/Elected Office (DO/EO/US) the following items and other information:			
<ol style="list-style-type: none"> <input checked="" type="checkbox"/> This is a FIRST submission of items concerning a filing under 35 U.S.C. 371. <input type="checkbox"/> This is a SECOND or SUBSEQUENT submission of items concerning a filing under 35 U.S.C. 371. <input type="checkbox"/> This express request to begin national examination procedures (35 U.S.C. 371(f)) at any time rather than delay examination until the expiration of the applicable time limit set in 35 U.S.C. 371(b) and PCT Articles 22 and 39(1). <input type="checkbox"/> A proper Demand for International Preliminary Examination was made by the 19th month from the earliest claimed priority date. <input checked="" type="checkbox"/> A copy of the International Application as filed (35 U.S.C. 371(c)(2)) <ol style="list-style-type: none"> <input checked="" type="checkbox"/> is transmitted herewith (required only if not transmitted by the International Bureau). <input type="checkbox"/> has been transmitted by the International Bureau. <input type="checkbox"/> is not required, as the application was filed in the United States Receiving Office (RO/US) <input checked="" type="checkbox"/> A translation of the International Application into English (35 U.S.C. 371(c)(2)). <input type="checkbox"/> Amendments to the claims of the International Application under PCT Article 19 (35 U.S.C. 371(c)(3)) <ol style="list-style-type: none"> <input type="checkbox"/> are transmitted herewith (required only if not transmitted by the International Bureau). <input type="checkbox"/> have been transmitted by the International Bureau. <input type="checkbox"/> have not been made; however, the time limit for making such amendments has NOT expired. <input type="checkbox"/> have not been made and will not be made. <input type="checkbox"/> A translation of the amendments to the claims under PCT Article 19 (35 U.S.C. 371(c)(3)). <input checked="" type="checkbox"/> An oath or declaration of the inventor(s) (35 U.S.C. 371(c)(4)). Unexecuted <input type="checkbox"/> A translation of the annexes to the International Preliminary Examination Report under PCT Article 36 (35 U.S.C. 371(c)(5)). 			
Items 11. to 16. below concern other document(s) or information included:			
<ol style="list-style-type: none"> <input checked="" type="checkbox"/> An Information Disclosure Statement under 37 CFR 1.97 and 1.98. <input type="checkbox"/> An assignment document for recording. A separate cover sheet in compliance with 37 CFR 3.28 and 3.31 is included. <input checked="" type="checkbox"/> A FIRST preliminary amendment. <input type="checkbox"/> A SECOND or SUBSEQUENT preliminary amendment. <input type="checkbox"/> A substitute specification. <input type="checkbox"/> A change of power of attorney and/or address letter. <input checked="" type="checkbox"/> Other items or information: Drawings (60 sheets); French International Preliminary Examination Report; Paper Copy of Sequence Listing in English 			

09/463705		INTERNATIONAL APPLICATION NO PCT/FR98/01593		ATTORNEY'S EXAMINATION NUMBER 146.1335	
17. <input checked="" type="checkbox"/> The following fees are submitted:				CALCULATIONS	
Basic National Fee (37 CFR 1.492(a)(1)-(5)):					
Search Report has been prepared by the EPO or JPO..... \$830.00				\$ 970.00	
International preliminary examination fee paid to USPTO (37 CFR 1.482)					
..... \$640.00					
No international preliminary examination fee paid to USPTO (37 CFR 1.482)					
but international search fee paid to USPTO (37 CFR 1.445(a)(2)).. \$710.00					
Neither international preliminary examination fee (37 CFR 1.482) nor					
international search fee (37 CFR 1.445(a)(2)) paid to USPTO..... \$950.00					
International preliminary examination fee paid to USPTO (37 CFR 1.482)					
and all claims satisfied provisions of PCT Article 33(2)-(4)..... \$90.00					
ENTER APPROPRIATE BASIC FEE AMOUNT =				\$ 970.00	
Surcharge of \$130.00 for furnishing the oath or declaration later than <input type="checkbox"/> 20 <input type="checkbox"/> 30					
months from the earliest claimed priority date (37 CFR 1.492(e)).					
Claims	Number Filed	Number Extra	Rate		
Total claims	41 -20 -	21	X \$ 18	\$ 378.00	
Independent Claims	21 -3 -	18	X \$ 78	\$ 1404.00	
Multiple dependent claims(s) (if applicable)			+ \$230.00	\$	
TOTAL OF ABOVE CALCULATIONS =				\$ 2752.00	
Reduction by 1/2 for filing by small entity, if applicable. Verified Small Entity statement				\$	
must also be filed. (Note 37 CFR 1.9, 1.27, 1.28).					
SUBTOTAL =				\$ 2752.00	
Processing fee of \$130.00 for furnishing the English translation later the <input type="checkbox"/> 20 <input type="checkbox"/> 30					
months from the earliest claimed priority date (37 CFR 1.492(f)).				+	
TOTAL NATIONAL FEE =				\$ 2752.00	
Fee for recording the enclosed assignment (37 CFR 1.21(h)). The assignment must be				\$	
accompanied by an appropriate cover sheet (37 CFR 3.28, 3.31). \$40.00 per property +					
TOTAL FEES ENCLOSED =				\$ 2752.00	
				Amount to be:	
				refunded \$	
				charged \$	

- a. ☒ A check in the amount of \$ 2752.00 to cover the above fees is enclosed.
- b. ☐ Please charge my Deposit Account No. _____ in the amount of \$ _____ to cover the above fees.
A duplicate copy of this sheet is enclosed.
- c. ☒ The Commissioner is hereby authorized to charge any additional fees which may be required, or credit any overpayment to Deposit Account No. 02-2275. A duplicate copy of this sheet is enclosed.

NOTE: Where an appropriate time limit under 37 CFR 1.494 or 1.495 has not been met, a petition to revive (37 CFR 1.137(a) or (b)) must be filed and granted to restore the application to pending status.

SEND ALL CORRESPONDENCE TO:

Bierman, Muserlian and Lucas
600 Third Avenue
New York, NY 10016

SIGNATURE

Charles A. Muserlian

NAME

19,683

REGISTRATION NUMBER

09/463705

428 Rec'd PCT/PTO 25 JAN 2000

Our Ref.: 146.1335

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Application of: :
PCT/FR98/01593 : PCT Date: July 21, 1998
FROMENTIN et al :
Serial No.: :
Filed: Concurrently Herewith :
For: BIOSYNTHESIS...AND THEIR USE :
600 Third Avenue
New York, NY 10016

PRELIMINARY AMENDMENT

Assistant Commissioner for Patents
Washington, D.C. 20231

Sir:

Please amend this application as follows:

IN THE SPECIFICATION:

Page 1, before line 1, insert

--This application is a 371 of PCT/FR98/01593 filed July
21, 1998.--

IN THE CLAIMS:

Claim 5, line 1, cancel "one of claims 1 to 4" and insert
--claim 1--.


Claim 12, line 1, cancel "one of claims 8 to 11" and insert
--claim 8--.

09/463705

REMARKS

The amendment is submitted to insert reference to the PCT application and to remove multiple dependency from the claims.

Respectfully submitted,
BIERMAN, MUSERLIAN AND LUCAS


Charles A. Muserlian, #19,683
Attorney for Applicant(s)
Tel. # (212) 661-8000

CAM:sd
Enclosure: Return Receipt Postcard

002220 50729450

Biosynthesis and transfer genes of 6-desoxyhexoses in
Saccharopolyspora erythraea and in *Streptomyces antibioticus*
and their use.

5 The present invention describes the genes involved in the biosynthesis and transfer of 6-desoxyhexoses in *Saccharopolyspora erythraea* and their use in the production of analogues of erythromycin by genetic manipulation.

10 Erythromycin A is a clinically important macrolide antibiotic produced by the gram-positive bacterium *Sac. erythraea*. The biosynthesis genes of erythromycin are organised in a cluster of *ery* genes which also includes the erythromycin auto-resistance gene *ermE*.

15 The *ery* cluster contains the three major genes *eryAI*, *eryAII* and *eryAIII* (locus *eryA*) coding for three polypeptides comprising the polyketide synthetase (called PKS) flanked by two regions comprising the genes involved in the subsequent stages of the conversion of the lactone nucleus (6-desoxyerythronolide B) into erythromycin A.

20 During the erythromycin A biosynthesis process represented in Figure 1, the biosynthesis of the 6-desoxyhexoses comprises all the enzymatic reactions leading from glucose-1-phosphate to the final activated sugar dTDP-L-mycarose or dTDP-D-desosamine. The dTDP-L-mycarose or dTDP-D-desosamine thus produced are then used as substrates for the transfer of the two desoxyhexoses onto the lactone nucleus. The formation of erythromycin requires the attachment of the mycarose via the hydroxyl in position C-3 of the lactone nucleus and the attachment of the desosamine via the hydroxyl in position C-5. All the *eryB* genes involved in the biosynthesis or the transfer of the mycarose and all the *eryC* genes involved in the biosynthesis or transfer of the desosamine have not yet been clearly identified.

35 The *ery* cluster with a length of 56 kb comprises 21 open reading frames (ORFs) the numbering of which has been established by Haydock et al. (1991) and Donadio et al. (1993). The *eryA* locus comprises the ORFs 10, 11 and 12.

005220 50429460

Early work on gene interruption or replacement in the left part of the ery cluster allowed a first identification of the eryCI gene (ORF1) (Dhillon et al., 1989), then of the eryBI gene (ORF2), of the eryH locus (ORFs 3, 4 and 5) the inactivation of which leads to the production of 6-desoxyerythronolide B, of a eryBII locus (ORFs 7 and 8) and the eryCII gene (Weber et al., 1990).

Among the enzymatic activities involved in the subsequent modifications of the lactone nucleus the eryF gene (ORF4) responsible for the hydroxylation in position C6 (Weber et al., 1991) and the eryK gene (ORF20) responsible for the hydroxylation in C12 (Stassi et al., 1993) have been identified. Moreover, the eryG gene (ORF6) responsible for the O-methylation of mycarose to cladinose (position 3"OH) has been identified (Weber et al., 1989). Erythromycin A is thus formed via erythromycin B or erythromycin C from erythromycin D according to the proposed diagram (Figure 1).

The functional characterization of the eryB and eryC genes situated on the right part of the ery cluster (ORFs 13 to 19) has not yet been precisely established, despite the fragmentary information reported in various journals (Donadio et al., 1993; Liu and Thorson, 1994; Katz and Donadio, 1995).

Due to the commercial interest in macrolide antibiotics, the obtaining of new derivatives, in particular of analogues of erythromycin having advantageous properties, is keenly sought. Modifications may be desired in the aglycone part (macrolactone) or/and in its secondary hydroxylation as well as in the sugar part (cladinose and/or desosamine) of erythromycin.

Current methods such as chemical modifications are difficult and limited vis-à-vis the type of product that can be obtained from erythromycin. For example, Sakakibara et al. (1984) review chemical modifications carried out from erythromycin A or B, both in the sugar part and the macrolactone.

Modifications of the macrolactone of erythromycin A by genetic manipulation of the microorganism *Sac. erythraea* have been described in the International Patent Application

00463705 032300

WO 93/13663 as well the obtaining of novel polyketide molecules by specific genetic alterations of the *eryA* locus of the chromosome coding for PKS. For example 7-hydroxyerythromycin A, 6-desoxy-7-hydroxyerythromycin A or 3-oxo-3-desoxy-5-desoaminy-erythronolide A have been obtained in this way.

The present invention relates to the functional characterization of ten genes of *Sac. erythraea* involved in the biosynthesis or the attachment of mycarose and of desosamine (*eryBII*, *eryCIII* and *eryCII* situated downstream from the *eryA* locus and *eryBIV*, *eryBV*, *eryCVI*, *eryBVI*, *eryCIV*, *eryCV* and *eryBVII* situated upstream), their use in the production of analogues of erythromycin as well as a preparation process for them.

Therefore a subject of the present invention is an isolated single- or double-stranded DNA sequence, represented in Figure 2 (direct and complementary sequence of SEQ ID No 1) corresponding to the *eryG-eryAIII* region of the cluster of erythromycin biosynthesis genes and a particular subject is an above DNA sequence comprising:

- the *eryBII* sequence corresponding to ORF7 (complementary sequence of SEQ ID No.1 from nucleotide 48 to nucleotide 1046) and coding for a dTDP-4-keto-L-6-desoxyhexose 2,3-reductase,
- the *eryCIII* sequence corresponding to ORF8 (complementary sequence of SEQ ID No.1 from nucleotide 1046 to nucleotide 2308) and coding for a desosaminyltransferase and
- the *eryCII* sequence corresponding to ORF9 (complementary sequence of SEQ ID No.1 from nucleotide 2322 to nucleotide 3404) and coding for a dTDP-4-keto-D-6-desoxyhexose 3,4-isomerase.

The above DNA sequence shown in Figure 2 is a genomic DNA sequence which can be obtained for example by sub-cloning restriction fragments of a genomic DNA fragment of *Sac. erythraea*, according to operating conditions a detailed description of which is given hereafter.

A more particular subject of the invention is an isolated DNA sequence represented in Figure 2 chosen from the

eryBII sequence corresponding to ORF7 (complementary sequence of SEQ ID No.1 from nucleotide 48 to nucleotide 1046), the eryCIII sequence corresponding to ORF8 (complementary sequence of SEQ ID No.1 from nucleotide 1046 to nucleotide 2308) or the eryCII sequence corresponding to ORF9 (complementary sequence of SEQ ID No.1 from nucleotide 2322 to nucleotide 3404) and the sequences which hybridize and/or have significant homologies with this sequence or fragments of it and having the same function.

A quite particular subject of the invention is the eryCIII isolated DNA sequence represented in Figure 2 corresponding to ORF8 (complementary sequence of SEQ ID No.1 from nucleotide 1046 to nucleotide 2308 = complementary sequence of SEQ ID No.4) and coding for a desosaminyl-transferase.

The eryBII sequence corresponding to ORF7 codes for a polypeptide having 333 amino acids (sequence of SEQ ID No.2), the eryCIII sequence corresponding to ORF8 codes for a polypeptide having 421 amino acids (sequence of SEQ ID No.5) and the eryCII sequence corresponding to ORF9 codes for a polypeptide having 361 amino acids (sequence of SEQ ID No.3).

The respective enzymatic activities indicated above has been demonstrated by the introduction of an internal deletion to the corresponding gene as illustrated hereafter in the experimental part.

Sequences which hybridize and having the same function are taken to include DNA sequences which hybridize with one of the above DNA sequences under standard conditions of high or medium stringency described by Sambrook et al. (1989) and which code for a protein having the same enzymatic function. The term same enzymatic function, is taken to mean a given enzymatic activity on substrates of a similar nature, for example a dTDP-6-desoxyhexose or a bare or glycosylated macrolactone. Conditions of high stringency comprise for example a hybridization at 65°C for 18 hours in a 5 x SSPE, 10 x Denhardt, 100 µg/ml DNAss, 1 % SDS solution, followed by 2 washings for 20 minutes with a 2 x SSC, 0.05 % SDS solution at 65°C followed by a last washing for 45 minutes in a

0.1 x SSC, 0.1 % SDS solution at 65°C. Conditions of medium stringency comprise for example a last washing for 20 minutes in a 0.2 x SSC, 0.1 % SDS solution at 65°C.

Sequences which display significant homologies and
5 having the same function are taken to include sequences having an identity of nucleotide sequence of at least 60 % with one of the above DNA sequences and which code for a protein having the same enzymatic function.

A subject of the invention is also a polypeptide coded
10 by one of the above DNA sequences and a particular subject is a polypeptide corresponding to an ORF represented in Figure 2, chosen from ORF7 (having the sequence of SEQ ID No.2), ORF8 (having the sequence of SEQ ID No.5) or ORF9 (having the sequence of SEQ ID No.3) and the analogues of this
15 polypeptide.

Analogues are taken to include the peptides having an amino acid sequence modified by substitution, deletion or addition of one or more amino acids inasmuch as these products retain the same enzymatic function. The modified
20 sequences can be prepared, for example, by using the site-directed mutagenesis technique known to a person skilled in the art.

A more particular subject of the invention is the polypeptide corresponding to ORF 8 represented in Figure 2
25 (having the sequence of SEQ ID No.5) and having a desosaminyltransferase activity, called EryCIII.

The invention describes an EryCIII recombinant protein of *Sac. erythraea* obtained by expression in a host cell according to known genetic engineering and cell culture
30 methods.

The obtaining of the purified recombinant protein has allowed confirmation of the characterization of the glycosyltransferase function associated with the product of the eryCIII gene in an in-vitro test which demonstrates the
35 transfer of the active sugar dTDP-D-desosamine onto the lactone nucleus.

A subject of the invention is also thymidine 5'-(tri-hydrogen diphosphate), P'-[3,4,6-tridesoxy-3-(dimethylamino)-

006220 50/29460

D-.xylo.-hexopyranosyl] ester (dTDP-D-desosamine) and the addition salts with the bases, a preparation example of which is described hereafter in the experimental part.

A subject of the invention is also an isolated DNA sequence represented in Figure 3 (sequence of SEQ ID No.6) corresponding to the *eryAI-eryK* region of the cluster of erythromycin biosynthesis genes and a particular subject is an above DNA sequence comprising:

- the *eryBIV* sequence corresponding to ORF13 (sequence of SEQ ID No.6 from nucleotide 242 to nucleotide 1207) and coding for a dTDP-4-ceto-L-6-desoxyhexose 4-reductase,
- the *eryBV* sequence corresponding to ORF14 (sequence of SEQ ID No.6 from nucleotide 1210 to nucleotide 2454) and coding for a mycarosyltransferase,
- the *eryCVI* sequence corresponding to ORF15 (sequence of SEQ ID No.6 from nucleotide 2510 to nucleotide 3220) and coding for a dTDP-D-6-desoxyhexose 3-N-methyltransferase,
- the *eryBVI* sequence corresponding to ORF16 (sequence of SEQ ID No.6 from nucleotide 3308 to nucleotide 4837) and coding for a dTDP-4-ceto-L-6-desoxyhexose 2,3-deshydratase,
- the *eryCIV* sequence corresponding to ORF17 (sequence of SEQ ID No.6 from nucleotide 4837 to nucleotide 6039) and coding for a dTDP-D-6-desoxyhexose 3,4-deshydratase,
- the *eryCV* sequence corresponding to ORF18 (sequence of SEQ ID No.6 from nucleotide 6080 to nucleotide 7546) and coding for a dTDP-D-4,6-didesoxyhexose 3,4-reductase and
- the *eryBVII* sequence corresponding to ORF19 (sequence of SEQ ID No.6 from nucleotide 7578 to nucleotide 8156) and coding for a dTDP-4-ceto-D-6-desoxyhexose 3,5 epimerase.

The above DNA sequence shown in Figure 3 is a genomic DNA sequence which can be obtained, for example, by sub-cloning restriction fragments of cosmids containing a genomic DNA library of *Sac. erythraea*, according to the operating conditions a detailed description of which is given hereafter.

A more particular subject of the invention is an isolated DNA sequence represented in Figure 3 chosen from the *eryBIV* sequence corresponding to ORF13 (sequence of SEQ ID

No.6 from nucleotide 242 to nucleotide 1207), the eryBV sequence corresponding to ORF14 (sequence of SEQ ID No.6 from nucleotide 1210 to nucleotide 2454), the eryCVI sequence corresponding to ORF15 (sequence of SEQ ID No.6 from nucleotide 2510 to nucleotide 3220), the eryBVI sequence corresponding to ORF16 (sequence of SEQ ID No.6 from nucleotide 3308 to nucleotide 4837), the eryCIV sequence corresponding to ORF17 (sequence of SEQ ID No.6 from nucleotide 4837 to nucleotide 6039), the eryCV sequence corresponding to ORF18 (sequence of SEQ ID No.6 from nucleotide 6080 to nucleotide 7546) or the eryBVII sequence corresponding to ORF19 (sequence of SEQ ID No.6 from nucleotide 7578 to nucleotide 8156) and the sequences which hybridize and/or display significant homologies with this sequence or fragments of the latter and having the same function.

A quite particular subject of the invention is the eryBV isolated DNA sequence represented in Figure 3 corresponding to ORF14 (sequence of SEQ ID No.6 from nucleotide 1210 to nucleotide 2454) and coding for a mycarosyltransferase.

The eryBIV sequence corresponding to ORF13 codes for a polypeptide having 322 amino acids (SEQ ID No.7), the eryBV sequence corresponding to ORF14 codes for a polypeptide having 415 amino acids (SEQ ID No.8), the eryCVI sequence corresponding to ORF15 codes for a polypeptide having 237 amino acids (SEQ ID No.9), the eryBVI sequence corresponding to ORF16 codes for a polypeptide having 510 amino acids (SEQ ID No.10), the eryCIV sequence corresponding to ORF17 codes for a polypeptide having 401 amino acids (SEQ ID No.14), the eryCV sequence corresponding to ORF18 codes for a polypeptide having 489 amino acids (SEQ ID No.11) and the eryBVII sequence corresponding to ORF19 codes for a polypeptide having 193 amino acids (SEQ ID No.12).

The respective enzymatic activities indicated above have been demonstrated by the introduction of an internal deletion to the corresponding gene as illustrated hereafter in the experimental part.

The DNA sequences which hybridize as well as the DNA

sequences which display significant homologies and having the same function have the same meaning as that indicated previously.

5 A subject of the invention is also a polypeptide coded by one of the above DNA sequences and a particular subject is a polypeptide corresponding to an ORF represented in Figure 3, chosen from ORF13 (having the sequence of SEQ ID No.7), ORF14 (having the sequence of SEQ ID No.8), ORF15 (having the sequence of SEQ ID No.9), ORF16 (having the sequence of SEQ
10 ID No.10), ORF17 (having the sequence of SEQ ID No.14), ORF18 (having the sequence of SEQ ID No.11) or ORF19 (having the sequence of SEQ ID No.12) and the analogues of this peptide.

The analogues of the polypeptide have the same meaning as that indicated previously.

15 A more particular subject of the invention is the polypeptide corresponding to ORF14 represented in Figure 3 (having the sequence of SEQ ID No.8) and having a mycarosyltransferase activity, called EryBV.

20 Knowledge of each eryB or eryC DNA sequence of the invention indicated above and shown in Figure 2 or Figure 3 allows the present invention to be reproduced for example by known methods of chemical synthesis or by screening a genomic library using oligonucleotide synthesis probes by known techniques of hybridization or by PCR amplification.

25 The polypeptides of the invention can be obtained by known methods, for example by chemical synthesis or by the recombinant DNA method by expression in a procaryotic or eucaryotic host cell.

30 Another subject of the invention relates to the use of at least one of the DNA sequences chosen from the sequences eryBII (complementary sequence of SEQ ID No.1 from nucleotide 48 to nucleotide 1046), eryCIII (complementary sequence of SEQ ID No.1 from nucleotide 1046 to nucleotide 2308) or eryCII (complementary sequence of SEQ ID No.1 from nucleotide
35 2322 to nucleotide 3404) represented in Figure 2, eryBIV (sequence of SEQ ID No.6 from nucleotide 242 to nucleotide 1207), eryBV (sequence of SEQ ID No.6 from nucleotide 1210 to nucleotide 2454), eryCVI (sequence of SEQ ID No.6 from

nucleotide 2510 to nucleotide 3220), eryBVI (sequence of SEQ ID No.6 from nucleotide 3308 to nucleotide 4837), eryCIV (sequence of SEQ ID No.6 from nucleotide 4837 to nucleotide 6039), eryCV (sequence of SEQ ID No.6 from nucleotide 6080 to nucleotide 7546) or eryBVII (sequence of SEQ ID No.6 from nucleotide 7578 to nucleotide 8156) represented in Figure 3, to synthesize hybrid secondary metabolites in *Sac. erythraea*.

Hybrid secondary metabolites are taken to mean either analogues of erythromycin, i.e. derivatives of erythromycin having one or more modifications affecting the sugar part and having an antibiotic activity, or precursors of erythromycin such as 6-desoxyerythronolide B or erythronolide B to which one or more sugar residues modified or not and not having an antibiotic activity are attached. The modified sugar residue can be, for example, 4-keto-L-mycarose.

The synthesis of hybrid secondary metabolites in *Sac. erythraea* by the use of eryB or eryC DNA sequences of the invention can be carried out, for example, by the inactivation of one or more eryB or eryC genes above and the introduction of one or more exogenous genes or their derivatives obtained for example by mutagenesis, having nucleotide sequences coding for the enzymes having the same function in strains producing other macrolides, for example tylosin, picromycin or methymycin. In particular, the introduction of exogenous genes can be carried out by the integration of a DNA sequence obtained according to the "DNA shuffling" method (Stemmer, 1994) or by the construction of a chimeric DNA sequence, for example from an eryB or eryC sequence of the invention intervening in the transfer of a sugar residue, for example the eryCIII or eryBV sequence, and homologous genes isolated from strains producing macrolides, for example *Streptomyces fradiae*, *Streptomyces olivaceus*, *Streptomyces venezuelae* or *Streptomyces antibioticus*.

The invention also relates to the use of at least one of the DNA sequences chosen from the sequences eryBII (complementary sequence of SEQ ID No.1 from nucleotide 48 to nucleotide 1046), eryCIII (complementary sequence of SEQ ID No.1 from nucleotide 1046 to nucleotide 2308) or eryCII

002220"50229450

(complementary sequence of SEQ ID No.1 from nucleotide 2322 to nucleotide 3404) represented in Figure 2, eryBIV (sequence of SEQ ID No.6 from nucleotide 242 to nucleotide 1207), eryBV (sequence of SEQ ID No.6 from nucleotide 1210 to nucleotide 2454), eryCVI (sequence of SEQ ID No.6 from nucleotide 2510 to nucleotide 3220), eryBVI (sequence of SEQ ID No.6 from nucleotide 3308 to nucleotide 4837), eryCIV (sequence of SEQ ID No.6 from nucleotide 4837 to nucleotide 6039), eryCV (sequence of SEQ ID No.6 from nucleotide 6080 to nucleotide 7546) or eryBVII (sequence of SEQ ID No.6 from nucleotide 7578 to nucleotide 8156) represented in Figure 3 or of a fragment of this sequence, as hybridization probes.

The eryB or eryC DNA sequences of the invention can be used to constitute hybridization probes of at least 19 nucleotides, allowing homologous genes to be isolated in strains producing macrolides by using the standard methods of hybridization of nucleic acids immobilized on filters or of PCR amplification, according to the conditions described by Sambrook et al. (1989).

The invention particularly relates to the use of the eryCIII DNA sequence represented in Figure 2 (complementary sequence of SEQ ID No.1 from nucleotide 1046 to nucleotide 2308 = complementary sequence of SEQ ID No.4) as a hybridization probe for isolating homologous genes responsible for the glycosylation of the macrolactone in a productive strain of macrolide.

The invention more particularly relates to the above use, in which the homologous genes are the oleandomycin biosynthesis genes in *S. antibioticus*.

The invention describes, by way of an example, the use of the sequence of the eryCIII gene as a hybridization probe for isolating homologous genes in a productive strain of oleandomycin. The eryCIII probe used has allowed the isolation of the *oleG1* and *oleG2* genes coding for glycosyltransferases in *S. antibioticus* involved in the transfer of desosamine and oleandroside onto the lactone nucleus.

The functional characterization of the *oleG1* and *oleG2*

genes has allowed the organisation of the right part of the cluster of oleandomycin biosynthesis genes in *S. antibioticus* to be defined.

- Therefore a subject of the invention is an isolated DNA sequence represented in Figure 22 (sequence of SEQ ID No.15) corresponding to a region of the cluster of oleandomycin biosynthesis genes comprising:
- the sequence corresponding to ORF *oleP1* from nucleotide 184 to nucleotide 1386,
 - the sequence corresponding to ORF *oleG1* from nucleotide 1437 to nucleotide 2714 coding for a glycosyltransferase activity,
 - the sequence corresponding to ORF *oleG2* from nucleotide 2722 to nucleotide 3999 coding for a glycosyltransferase activity,
 - the sequence corresponding to ORF *oleM* from nucleotide 3992 to nucleotide 4720 (= sequence of SEQ ID No.20) and
 - the sequence corresponding to ORF *oleY* from nucleotide 4810 to nucleotide 5967.

- The above DNA sequence shown in Figure 22 (sequence of SEQ ID No.15) is a genomic DNA sequence which can be obtained for example from a cosmid covering the right part of the cluster of oleandomycin biosynthesis genes by hybridization with an *eryCIII* probe, according to the operating conditions a detailed description of which is given hereafter.

- A more particular subject of the invention is an isolated DNA sequence represented in Figure 22 chosen from the sequence corresponding to ORF *oleG1* (sequence of SEQ ID No.15 from nucleotide 1437 to nucleotide 2714 coding for a glycosyltransferase activity and the sequence corresponding to ORF *oleG2* (sequence of SEQ ID No.15 from nucleotide 2722 to nucleotide 3999) coding for a glycosyltransferase activity.

- A quite particular subject of the invention is an isolated DNA sequence above corresponding to ORF *oleG1* (sequence of SEQ ID No.15 from nucleotide 1437 to nucleotide 2714) coding for a desosaminyltransferase activity, as well as an isolated DNA sequence above corresponding to ORF *oleG2*

005220" 50' 59' 50"

(sequence of SEQ ID No.15 from nucleotide 2722 to nucleotide 3999) coding for an oleandrosyltransferase activity.

The sequence corresponding to ORF oleG1 codes for a polypeptide having 426 amino acids (sequence of SEQ ID No.17) and the sequence corresponding to ORF oleG2 codes for a polypeptide having 426 amino acids (sequence of SEQ ID No.18).

The respective enzymatic activities indicated above have been demonstrated by alteration of the corresponding gene as illustrated hereafter in the experimental part.

10 A subject of the invention is also the polypeptide coded
by the DNA sequence corresponding to ORF *oleG1* and having a
desosaminyltransferase activity (sequence of SEQ ID No.17)
and the polypeptide coded by the DNA sequence corresponding
15 to ORF *oleG2* and having an oleandrosyltransferase activity
(sequence of SEQ ID No.18).

The polypeptides above called OleG1 and OleG2 respectively can be obtained by the known methods indicated above.

20 A subject of the invention is also a process for the preparation of hybrid secondary metabolites in *Sac. erythraea* in which:

in which:

- a DNA sequence is isolated containing at least one eryB sequence or an eryC sequence of the cluster of biosynthesis genes of erythromycin represented in Figure 2 (complementary sequence of SEQ ID No.1) or in Figure 3 (sequence of SEQ ID No.6),

No.6),
- a modification is created in said sequence and an altered sequence is obtained,

sequence is obtained,
30 - the altered sequence is integrated into the chromosome of
the host strain and a modified strain is obtained,

- the modified strain is cultured under conditions which allow the formation of the hybrid secondary metabolite and
- the hybrid secondary metabolite is isolated.

35 The modification of the DNA sequence can be carried out for example by an addition and/or by a deletion of DNA sequences of at least one nucleotide, in an *eryB* or *eryC* sequence of the invention which codes for one of the

corresponding enzymes indicated above.

The integration of the altered sequence into the host strain can be carried out for example by the homologous recombination method which can be carried out according to the diagram shown in Figure 4 and leads to the generation of chromosomal mutants of *Sac. erythraea* strains which are then cultured according to the known general methods of cell culture.

A particular subject of the invention is the above process in which the DNA sequence codes for one of the enzymes chosen from a

- dTDP-4-keto-D-6-desoxyhexose 2,3-reductase,
- desosaminyltransferase,
- dTDP-4-keto-D-6-desoxyhexose 3,4-isomerase,
- dTDP-4-keto-L-6-desoxyhexose 4-reductase,
- mycarosyltransferase,
- dTDP-D-6-desoxyhexose 3-N-methyltransferase,
- dTDP-4-keto-L-6-desoxyhexose 2,3-deshydratase,
- dTDP-D-6-desoxyhexose 3,4-deshydratase,
- dTDP-D-4,6-didesoxyhexose 3,4-reductase or
- dTDP-4-keto-D-6-desoxyhexose 3,5 epimerase.

A more particular subject of the invention is the above process in which the alteration of the sequence results in the inactivation of at least one of the enzymes indicated above.

The inactivation of at least one of the enzymes is demonstrated, on the one hand by the absence of production of erythromycin, on the other hand by the accumulation of precursors of erythromycin such as 6-desoxyerythronolide B, erythronolide B or 3- α -mycarosyl erythronolide B and/or the accumulation of hybrid secondary metabolites as defined previously in the supernatants of cultures of the corresponding modified strains.

The invention quite particularly relates to the above process in which the inactivated enzyme is a dTDP-4-keto-L-6-desoxyhexose 4-reductase or in which the inactivated enzyme is a dTDP-D-6-desoxyhexose 3,4-deshydratase or in which the enzyme is a mycarosyltransferase or in which the inactivated

The invention also relates to the above process in which the isolated hybrid secondary metabolite is an analogue of erythromycin chosen from 4"-keto-erythromycin, 4'-hydroxy-erythromycin or 3"-C desmethyl-2",3"-ene-erythromycin or in which the isolated hybrid secondary metabolite is desosaminyl erythronolide B.

The invention also relates to a modified strain of *Saccharomyces erythraea* in which at least one of the enzymes is chosen from a

- The invention particularly relates to the modified strain of *Sac. erythraea* BII92 in which a dTDP-4-keto-L-6-desoxyhexose 2,3-reductase is inactivated and producing 3"-C desmethyl-2",3"-ene-erythromycin C, the modified strain BIV87 of *Sac. erythraea* in which a dTDP-4-keto-L-6-desoxyhexose 4-reductase is inactivated and producing 4"-keto-erythromycin, the modified strain CIV89 of *Sac. erythraea* in which a dTDP-D-6-desoxyhexose 3,4-deshydratase is inactivated and producing 4'-hydroxyerythromycin D as well as the modified strain BV88 of *Sac. erythraea* in which a mycarosyltransferase is inactivated and producing desoaminy l erythronolide B. Detailed constructions of the above strains are given

hereafter in the experimental part.

The invention also relates to a process for the preparation of precursors of oleandomycin in *S. antibioticus* in which

- 5 - an alteration is created in the sequence of the gene chosen from the DNA sequence corresponding to ORF *oleG1* (sequence of SEQ ID No.15 from nucleotide 1437 to nucleotide 2714) and the DNA sequence corresponding to ORF *oleG2* (sequence of
- 10 SEQ ID No.15 from nucleotide 2722 to nucleotide 3999) in the chromosome of a host strain and a modified strain is obtained,
- the modified strain is cultured under conditions allowing the accumulation of the precursors of oleandomycin and
- 15 - these precursors are isolated.

The alteration of the DNA sequence can be carried out for example by interruption of the target gene in the strain *S. antibioticus*, for example by integration of a plasmid by the homologous recombinant method and leads to the generation

20 of chromosomal mutants of the wild strain.

The invention particularly relates to an above process in which the alteration is created in the DNA sequence corresponding to ORF *oleG1* (sequence of SEQ ID No.15 from nucleotide 1437 to nucleotide 2714) and the result of which

25 is at least the elimination of the desoaminyltransferase activity and the accumulation of the precursor of oleandomycin 8,8a-desoxyoleandolide.

The accumulation of a non-glycosylated precursor of oleandomycin 8,8a-desoxyoleandolide observed by interruption

30 of the *oleG1* gene is due to a transcriptional polar effect which inactivates the *oleG2* gene.

An example of implementation of the above process is given hereafter in the experimental part.

35 **Materials and general methods.**

1. Bacterial strains, plasmids and growth conditions.

The *Sac. erythraea* strain used for the realisation of the invention is a spontaneous phenotypic variant called "red

5

10

15

20

25

30

35

The general methods of molecular biology used are described by Sambrook et al., 1989.

Reagents of commercial origin were used including the restriction enzymes (New England Biolabs and Boehringer Mannheim), the Klenow fragment of DNA polymerase I (Boehringer Mannheim). The "DNA ligation system" kit (Amersham) was used to carry out the ligations and the Plasmid Midi kit (Quiagen) or RPM kit (Bio101 Inc.) for purifying the plasmidic DNA.

The preparation of the DNA of the λ bacteriophage was carried out according to Ausubel et al. (1995) and the isolation of the chromosomal DNA of *Sac. erythraea* according to Hopwood et al. (1985).

The transformation of *S. lividans* and isolation of the plasmids was carried out according to Hopwood et al. (1985).
3. Preparation of erythronolide B and 3- α -mycarosyl erythronolide B.

Erythronolide B and 3- α -mycarosyl erythronolide B were purified from culture extracts of the mutant *eryCI* (clone WHB2221 described by Dhillon et al., 1989) by chromatography on aminopropyl gel (LichroprepNH2 25-40 μ , Merck) with an elution gradient by successive butyl chloride/ methylene chloride mixtures (100:0, 80:20, 50:50 and 20:80) followed by a linear elution gradient by the butyl chloride /methanol mixture varying from 99:1 to 90:10. The fractions containing the expected products are brought to dryness under vacuum then analyzed by thin-layer chromatography (TLC). Erythronolide B is then crystallized from an ethyl acetate/hexane mixture then recrystallized from ethanol. 3- α -mycarosyl erythronolide B is crystallized twice from an ethyl acetate/hexane mixture.

Mentioned media.

1. R2T2:

For 1 litre of aqueous solution: sucrose 103 g; K₂SO₄ 0.25g; yeast extract 6.5 g; tryptone 5.0 g; bactoagar 22.0 g; distilled water sqf 860 ml. The solution is sterilized by autoclaving for 30 minutes at 120°C. At the time of use, the following sterile solutions are added: 20 ml of glucose at 50

09463705.022300

2. R2T20:

3. M1-102 (Kaneda et al., 1962):

36 mg; tap water 1000 ml; final pH adjusted to 7.0 to 7.2 with KOH. The solution is sterilized by autoclaving for 30 minutes at 120°C.

4. R2YE(R5) (Hopwood et al., 1985):

15 $\text{MgCl}_2 \cdot 6\text{H}_2\text{O}$ 10.12 g; casaminoacids 0.1 g; solution of trace
elements" 2 ml; yeast extract 5 g; TES 5.72 g; bactoagar 15
g; distilled water sqf 940 ml. The solution is sterilized by
autoclaving for 30 minutes at 120°C . At the time of use, the
following sterile solutions are added: 10 ml of 0.5 % KH_2PO_4 ;
20 20 ml of CaCl_2 1M; 15 ml of L-proline at 20 %; 20 ml of
glucose at 50 % and 1 ml of CuCl_2 10mM.

5. 2 x TY:

25 6. PT buffer:

7. Sucrose-succinate (Caffrey et al., 1992):

35 5 mM; potassium nitrate 100 mM; solution of "trace elements"
2 ml. The solution is sterilized by autoclaving for 30

The attached figures illustrate certain aspects of the

invention.

Figure 1 represents the biosynthesis route for erythromycin A.

Figure 2 represents the nucleotide sequence (direct and complementary sequence of SEQ ID No.1) of the *eryG-eryAIII* region of the cluster of biosynthesis genes of erythromycin comprising ORFs 7, 8 and 9 and their deduced protein sequences.

Figure 3 represents the nucleotide sequence (sequence of SEQ ID No.6) of the *eryAI-eryK* region of the cluster of biosynthesis genes of erythromycin comprising ORFs 13 to 19 and their deduced protein sequences.

Figure 4 represents the diagram for gene substitution by homologous recombination.

Figure 5A represents the organisation of the left part of the cluster of biosynthesis genes of erythromycin in *Sac. erythraea* the ORFs 1 to 9 of which are indicated by arrows as well as a restriction map of the plasmids pK62, pBCK1, pKB22, pBK44, pBIISB, pEco2 and pK23, generated from genomic clone λ SE5.5. (Abbreviations of the restriction enzymes: B, *Bam*HI; Bc, *Bcl*II; Bg, *Bgl*III; E, *Eco*RI; K, *Kpn*I; M, *Mlu*I; P, *Pst*I; S, *Sac*I; Sa, *Sal*I.)

Figure 5B represents the organisation of the right part of the cluster of the biosynthesis genes of erythromycin in *Sac. erythraea* the ORFs 13 to 21 of which are indicated by arrows as well as a restriction map of the plasmids pBK6-12, pCN9, pNCO28, pNB49, pNCO62, pPSP4, pNCO62X and pBAB18. (Abbreviations of the restriction enzymes: B, *Bam*HI; Ba, *Bal*I; Bc, *Bcl*II; C, *Cla*I; E, *Eco*RI; K, *Kpn*I; N, *Nco*I; Ns, *Nsi*I; P, *Pst*I; Pv, *Pvu*II; S, *Sac*I; Sc, *Sca*I; Sh, *Sph*I; Sp, *Spe*I; X, *Xba*I; Xh, *Xho*I).

Figure 6A represents the construction diagram for the plasmid pBIIA.

Figure 6B represents a restriction map of the plasmid pUWL218.

Figure 6C represents a restriction map of the plasmid pBIIA.

Figure 7A represents the construction diagram of the

the mutant strains BII92, CIII68, CII62, BIV87, BV88, CIV89 and CV90, compared with the "red variant" wild strain marked Wt. For each mutant, the restriction enzyme used is indicated below each blot and the size of the bands detected before each blot is estimated relative to the molecular weight markers λ -HindIII and λ -BstEII (not detectable by auto-radiography).

Figure 16 represents the analysis by PCR of the mutant strains BII91, CIII68, CII62, BIV87, BV88, CIV89 and CV90, compared with the "red variant" wild strain marked Wt and with the plasmids pBIIA, pCIIIA, pCIIA, pBIVA, pBVA, PCIVA and pCVA used respectively in order to obtain the mutant by homologous recombination. The sizes of the bands detected by ethidium bromide staining are estimated relative to the molecular weight markers Φ X174-HaeIII or λ -BstEII. Figure 17 represents the analysis by TLC of the metabolites produced by the mutant strains BII92, CIII68, CII62, BIV87, BV88, CIV89 and CV90, compared with the standard products erythromycin A (Er A), erythronolide B (EB) and 3- α -mycarosyl erythronolide B (MEB).

Figure 18 represents the analysis by SDS-PAGE of the purification of the EryCIII protein successively after extraction with urea 7M (line 2), Q Sepharose chromatography (line 3), Superdex chromatography (line 4), Q source chromatography (line 6) with standard molecular weight markers (lines 1 and 5);

Figure 19 represents the analysis by TLC of the biological activity test of the EryCIII protein, by incubation with d-TDP-D-desosamine (line 2) or with d-TDP-D-desosamine and 3- α -mycarosyl erythronolide B (MEB) (line 3) compared with the MEB control (line 1) and with the erythromycin A control (line 4). The dotted lines mark the zones showing an antibiotic activity using an autobiogram on *B. pumilus*.

Figure 20 represents the localization of the six cosmids (cosAB35, cosAB76, cosAB87, cosAB67, cosAB63 and cosAB61) covering all of the cluster of the oleandomycin biosynthesis genes. The restriction fragments *Bam*HI (marked B)

hybridizing with the probes marked str M, D, E and the *Bam*HI fragments (3.5 kb and 2.7 kb) hybridizing with the probe *eryCIII* are shown.

Figure 21 represents the organisation of the right part of the cluster of the oleandomycin biosynthesis genes in *S. antibioticus* the different ORFs of which (marked *oleP1*, *oleG1*, *oleG2*, *oleM*, *oleY*, *oleP* and *oleB*) are indicated by arrows as well as a restriction map of the plasmid pCO35-S and the insert of the plasmid pCO3 generated from pCO35-S. The double arrow indicates the insert corresponding to the sequence of Figure 22 (abbreviations of the restriction enzymes: B, *Bam*HI; Bg, *Bgl*II; K, *Kpn*I; S, *Sac*I; Sh, *Sph*I; the star indicates that it is not a unique site).

Figure 22 represents the nucleotide sequence (sequence of SEQ ID No.15) of the region covering the *oleP1*, *oleG1*, *oleG2*, *oleM* and *oleY* oleandomycin biosynthesis genes and their deduced protein sequences.

EXAMPLE 1: cloning and sequencing the *eryG-eryAIII* region of the cluster of biosynthesis genes of erythromycin.

A genomic DNA fragment of *Sac. erythraea* NRRL 2338 having > 20 kb downstream from the *ermE* gene covering in particular ORFs 3 to 9 and corresponding to the clone λ SE5.5 as well as the nucleotide sequence of a 4.5 kb fragment corresponding to the region of the *ery* cluster comprised between 3.7 kb and 8.0 kb from the 3' end of the *ermE* gene and comprising ORFs 3, 4, 5 and 6 were described by Haydock et al. (1991).

Taking account of the restriction map shown by Haydock et al. (1991), sub-clones were derived from clone λ SE5.5 by sub-cloning restriction fragments in pUC19. The plasmids pKB22, pBK44, pBIISB and pEco2 were thus generated according to Figure 5A in the following fashion:

From the DNA of clone λ SE5.5 digested by the *Kpn*I restriction enzyme, plasmids pK62 and pK66 were directly constructed by sub-cloning of the *Kpn*I fragment of 5.8 kb in pUC19, the plasmid pK66 corresponding to the same *Kpn*I fragment sub-cloned with an reversed orientation of the insert relative to the vector. The plasmid pKB22 containing

an insert of 2.9 kb was then derived from plasmid pK66 by excision of the *Bam*HI-*Bgl*III fragment (2.9 kb) covering ORF8 as well as a part of ORFs 7 and 9 by digestion with the *Bam*HI and *Bgl*III restriction enzymes. In the same way, the plasmid pKB44 containing a insert of 2.9 kb was obtained from the plasmid pK62 by excision of the *Bam*HI-*Bgl*III fragment (2.9 kb) covering the *eryG* gene corresponding to ORFs 5 and 6 and the *eryF* gene corresponding to ORF4.

The plasmid pBIISB was derived from the plasmid pBK44 by sub-cloning in pUC19 of the *Sal*I fragment of 600 bp obtained from the plasmid pBK44 digested by the *Sal*I restriction enzyme (Figure 5A).

From the DNA of the clone λ SE5.5 digested by the *Eco*RI restriction enzyme, the plasmid pEco2 was directly constructed by sub-cloning of the *Eco*RI fragment (2.2 kb) in pUC19.

The sub-clones pKB22, pBK44, pBIISB and pEco2 thus obtained were then sequenced. Analysis was carried out on the plasmidic DNA samples, purified beforehand on a Quiagen 100 column (Quiagen), on the ABI prism 377 automatic sequencer. The sequencing reactions were carried out by the Sanger method (1977) using conventional M13 primers or synthetic primers and fluorescent triphosphate dideoxynucleosides and Taq FS polymerase (Perkin Elmer) in the presence of 5 % of dimethylsulphoxide, the synthetic primers used having the following sequences:

C3R2	TCCTCGATGGAGACCTGCC	(SEQ ID No.22)
B2R1	GAGACCATGCCCAGGGAGT	(SEQ ID No.23)
C3S2	TCTGGGAGCCGCTCACCTT	(SEQ ID No.24)
C2R1	GACGAGGCCGAAGAGGTGG	(SEQ ID No.25)
C2S	GCACACCGGAATGGATGCG	(SEQ ID No.26)
fullC3S	CCGTCGAGCTCTGAGGTAA	(SEQ ID No.27)
fullC3R	GCCCGAGCCGCACGTGCGT	(SEQ ID No.28) and
C4	TGCACGCGCTGCTGCCGACC	(SEQ ID No.29).

Assembling of the sequence data was carried out with the Autoassembler™ software package (Applied Biosystem). The sequences were analyzed using the GCG set of software (Devereux 1984).

The nucleotide sequences obtained allowed the nucleotide sequence of 3412 bp of Figure 2 to be established (direct and complementary sequence of SEQ ID No.1) in which three ORFs (7, 8 and 9) were identified respectively from nucleotide 8957 to nucleotide 7959, from nucleotide 10219 to nucleotide 8957 and from nucleotide 11315 to nucleotide 10233 (numbered in Figure 2 from the *Bam*HI site situated at the 5' end of the *ermE* gene) (respectively complementary sequence of SEQ ID No.1 from nucleotide 48 to nucleotide 1046, from nucleotide 1046 to nucleotide 2308 and from nucleotide 2322 to nucleotide 3404) and corresponding respectively to the *eryBII*, *eryCIII* and *eryCII* genes according to Liu and Thorson (1994) the functional characterizations of which had not yet been identified. The three ORFs 7, 8 and 9 have the same orientation, reading being carried out from the 3' region of the *eryAIII* gene.

Specimens of *E. coli* XL1-blue containing the coding region for the above ORFs were deposited at the Collection Nationale de Cultures de Microorganismes (CNCM) INSTITUT PASTEUR, 25, Rue du Docteur Roux 75724 PARIS CEDEX 15 FRANCE, on the 16 July 1997:

- the plasmid pK62 comprising the coding sequence for ORF7, ORF8 and part of ORF9 under the number I-1897,
- the plasmid p*Eco*2 comprising the coding sequence for ORF9 and part of ORF8 under the number I-1899.

EXAMPLE 2: construction of the plasmid pBIIA.

An integration plasmid, called **pBIIA** and carrying a deletion in the *eryBII* gene coding for ORF7, was constructed according to the diagram in Figure 6A.

The *Bcl*II-*Bam*HI fragment of 598 bp was deleted in the plasmid pK62 obtained in Example 1 by digestion with the *Bcl*II and *Bam*HI enzymes. The resulting plasmid pBCK1 was then digested with the *Mlu*I and *Bgl*III restriction enzymes so as to delete a 853 bp fragment inside ORF7 from nucleotide 8011 to nucleotide 8863 of the sequence of Figure 2. After filling in the ends using the Klenow fragment of DNA polymerase I, the plasmid containing the deletion was religated and transformed in *E. coli* XL1-blue. From the plasmid p19BIIA

thus generated, the *KpnI*-*HindIII* fragment (4.3 kb) which carries the deletion was sub-cloned in the plasmid pUWL218 (Figure 6B). The presence of the 853 bp deletion from nucleotide 8011 to nucleotide 8863 in the plasmid pBIIA thus

5 generated (Figure 6C) was confirmed by sequencing. The plasmid pBIIA was then transferred into the *E. coli* DH5 α MRC strain, then used to transform *Sac. erythraea*.

EXAMPLE 3: construction of a *Sac. erythraea* ery BIIA strain (BII92).

10 The construction of a *Sac. erythraea* strain in which the eryBII gene carried an internal deletion such as that introduced into the plasmid pBIIA prepared in Example 2 and the integration process were carried out in the following fashion:

15 The preparation of the protoplasts was carried out according to the method described by Weber and Losick (1988), using PEG 3350 (Sigma) instead of PEG 1000 and a modified P buffer (called PT) containing MgCl₂·6H₂O 28 mM and without PO₄H₂K instead of the P, L or T buffers described, according

20 to the following operating conditions:

The cells (at least 10⁸ spores) of *Sac. erythraea* "red variant" (a sample of which was deposited at the Collection Nationale de Cultures de Microorganismes (CNCM) INSTITUT PASTEUR, 25, Rue du Docteur Roux 75724 PARIS CEDEX 15 FRANCE,

25 on the 16 July 1997 under the number I-1902) were grown in 50 ml of TBS medium for 3 to 5 days at 30°C, then washed with 10.3 % sucrose. The cells were resuspended in 50 ml of PT buffer containing 2 to 5 mg/ml of lysozyme (Sigma), then incubated at 30°C for 1 to 2 hours, with desegregation the

30 mycelium mass every 15 minutes until transformation of at least 50 % of the mycelium to protoplasts. The protoplasts were washed with 50 ml of PT buffer, resuspended in 12.5 to 25 ml of the same buffer, slowly frozen then stored at -80°C by 200 μ l aliquots.

35 For the transformation, one aliquot was thawed and 50 μ l were removed then transferred into a 15 ml tube. One to 10 μ g of plasmidic DNA pBIIA, prepared in Example 2 from the *E. coli* DH5 α MRC strain were dissolved in 5 to 10 μ l of TE buffer

002220 50/22450

(Tris HCl 10 mM pH 7.5, EDTA 1 mM) then deposited on the wall of the inclined tube to which was then added 0.5 ml of a solution of PEG 3350 in PT buffer prepared extemporaneously from an aqueous solution at 50 % which is diluted to one half in the 2 x PT buffer. After dilution with 3 to 5 ml of PT buffer then centrifugation at 2500 rpm for 15 mn, the pellet was dissociated in 0.5 ml of PT buffer and the suspension of transformed protoplasts thus obtained is immediately distributed onto 2 or 3 very dry R2T2 dishes (3 hours under a laminar flow hood). The dishes were then incubated at 32°C for 16 to 24 hours until the regeneration film of the protoplasts appears. From a stock solution of thiostrepton (Sigma) at 50 mg/ml in DMSO, an appropriate quantity was diluted in 0.5 to 1 ml of water then spread on the dishes so as to obtain a final concentration of 20 µg of thiostrepton/ml of gelose. After complete absorption of the antibiotic, the dishes were incubated at 32°C for 3 to 4 days, which allows visualisation of the transformants. The dishes were incubated for several more days until the spores were completely developed.

Selection of the integrants corresponding to the first recombination event (Figure 4) was carried out by replication of the sporulated dishes using velvet or by spreading a suspension of the spores onto R2T2 dishes containing thiostrepton then incubation at 32°C, which allows clones of potential integrants to grow.

For the selection of clones which have undergone a second recombination event (Figure 4), 5 to 10 thiostrepton-resistant clones obtained above were cultured in 8 ml of TSB liquid medium at 30°C for 3 to 4 days. 50 to 100 µl were removed and recultured under the same conditions. After 4 successive cycles of dilution and culture intended to encourage the loss of the thiostrepton resistance marker, protoplasts were prepared from the cells as indicated above, so as to expel the plasmid. The protoplasts were then spread onto R2T2 dishes so as to obtain individualized colonies the sensitivity to thiostrepton of which was determined by replication on R2T2 dishes containing thiostrepton.

Depending on the position of the second recombination event relative to the deletion site (Figure 4), it can be expected that the phenotype of the colonies which are sensitive to thiostrepton are of wild type or of mutated type carrying the deletion.

Among the colonies which are sensitive to thiostrepton, selection of the mutants having the *ery^r* phenotype was carried out by record of bacterial sensitivity to antibiotics on the *B. pumilus* ATCC 14884 strain which is sensitive to erythromycin. The *B. pumilus* strain was used as an indicator strain for evaluating the production of erythromycin in biological tests by record of bacterial sensitivity to antibiotics. The colonies were spread using a platinum loop onto R2T2 dishes, then incubated for 3 to 4 days at 30°C. The agar areas where the mutant has grown at confluency are then removed with a punch then placed on A-Merck dishes covered with a 4 ml overlayer of 0.5 x A Merck (Antibiotic agar No.1 Merck) inoculated with a suspension of *B. pumilus* spores, then incubated overnight at 37°C.

The presence of the expected deletion in the chromosome of the mutant (deletion of 853 bp from nucleotide 8011 to nucleotide 8863 in Figure 2) was then confirmed by genomic analysis using Southern blot as well as by PCR in the following fashion:

For the analysis by Southern blot, the transfer of genomic DNA, digested beforehand with the appropriate restriction enzyme, on GeenscreenPlus (Dupont NEN) membranes was carried out in 0.4 M NaOH according to Ausubel et al. (1995). The hybridizations were carried out using as a probe the oligonucleotide labelled at its 5' end using [γ^{32} P]ATP (Amersham) and the polynucleotide kinase (Boehringer Mannheim) according to Sambrook et al. (1989), having the following sequence:

B2-S TTGGCGAAGTCGACCAGGTC (SEQ ID No.30)

corresponding to the DNA region at the start of the *eryG* gene situated from position 4118 to position 4137 of the sequence deposited in the EMBL base under the reference X60379 and described by Haydock et al. (1991). The hybridizations were

00463705 000000

carried out with a rapid hybridization buffer (Amersham) and the following washing conditions: 2 x 5 mn, 2 x SSC, 20°C; 30 mn, 2 x SSC, 65°C; 30 mn, 0.1 x SSC, 20°C.

By Southern hybridization on the genomic DNA isolated according to Hopwood et al. (1985) then digested by the *Kpn*I restriction enzyme, a band of 5.8 kb from the "red variant" wild strain and a band of 4.9 kb from the mutant BII92 were detected. The results shown in Figure 15 indicate the presence in the mutant of a deletion of approximately 900 bp in this region of the chromosome.

For the analysis by PCR, a 100 µl sample of a 3-day culture in TSB medium was centrifuged. The pellet obtained was resuspended in 10 µl of TSB medium, then used for amplification in the genAmp PCR system 9600 apparatus (Perkin Elmer Cetus). After heating the sample for 3 minutes at 94°C, the following amplification conditions were used: 94°C, 1 mn; 55°C, 1 mn; 72°C, 3 mn; 30 cycles; Ampli Taq polymerase (Perkin Elmer) in the presence of 10 % dimethylsulphoxide (v/v) followed by an extension of 3 mn at 72°C. The amplification was carried out using the oligonucleotide B2S above and the oligonucleotide having the following sequence: B2-R GCCGCTCGGCACGGTGAACCTCA (SEQ ID No.31) corresponding to the sequence of the complementary strand of the DNA region situated from position 8873 to position 8892 of the sequence of Figure 2 to which three nucleotides were added at the 5' end and which allows framing by PCR amplification of the region carrying the internal deletion at ORF7.

The analysis by PCR amplification on whole cells allowed a band of approximately 1 kb in the wild strain and a band of 0.16 kb in the mutant BII92 to be detected in an identical fashion to the signal obtained with the plasmid pBIIA. The results shown in Figure 16 confirm that the deletion of approximately 900 bp detected by the Southern analysis is identical to that carried by the plasmid pBIIA (853 bp).

The recombinant strain thus obtained, designated BII92, was then cultivated in order to identify the metabolites produced by the strain.

00463705 002300
00463705 002300

EXAMPLE 4: fermentation of the BII92 strain and identification of the secondary metabolites produced.

Extracts of the broth culture medium of the strain were analyzed by thin-layer chromatography (TLC) with erythromycin A, erythronolide B and 3- α -mycarosyl erythronolide B as standards.

The BII92 strain was cultured in a 50 ml erlenmeyer flask under conditions allowing an optimal production of erythromycin A and its derivatives which consist of carrying out a cell preculture at 28°C for 48 hours in EP1 medium (Solulys L-Corn steep liquor (Roquette frères) 5 g/l; defatted soya bean flour (Cargill) 10 g/l; CO₃Ca 2 g/l; NaCl 5 g/l; pH = 6.8; glucose sqf 15 g/l added after autoclaving), then a culture for 72 hours after dilution to 7 % v/v with EP2medium (defatted soya bean flour 10 g/l; CO₃Ca 0.2 g/l; Cl₂Co-6H₂O 1 mg/l; pH = 6,8-7.0; glucose sqf 20 g/l added after autoclaving).

The culture supernatant was then extracted at pH 9-10 with ethyl acetate. The organic phases were dried over MgSO₄, brought to dryness under reduced pressure then analyzed by TLC on 60 F254 silica gel (Merck) [dichloromethane/methanol (90:10, v/v) or isopropyl ether /methanol/NH₄OH at 25 % (75:35:2, v/v)]. Alternatively, the analysis was carried out by TLC on bonded silica gel plates of NH₂ F254 type (Merk) [butyl chloride/methanol (90:10, v/v)].

The chemical revelation of the plates was carried out by pulverization of a solution of p-anisaldehyde- 98 %sulphuric acid -ethanol (1:1:9, v/v), followed by heating for a few minutes at 80°C. The potential antibiotic activities were analyzed by direct bioautography of the TLC plates on agar seeded with *B. pumilus* ATCC 14884.

The results obtained by chemical revelation (Figure 17) shows that the BII92 strain preferentially accumulates erythronolide B as expected of an eryB mutant.

Minor metabolites of low mobility showing an antibiotic activity were also detected. These metabolites were extracted with ethyl acetate and identified by reversed-phase high-performance liquid chromatography (RP-HPLC) coupled with

mass spectrometry. The RP-HPLC was carried out on a column (250 x 4.6mm) of Kromasil C18 5 μ using an acetonitrile/methanol/ammonium acetate 0.065 M pH 6.7 mixture (350:150:500, v/v) as the mobile phase, on a Waters chromatograph equipped with a Finningan TSQ 7000 mass spectrometer.

Alongside traces of erythromycin A, B, C and D, 4 minor metabolites called M1 to M4 were detected:

- M1 produces a parent peak at m/z 704 and fragmentation products at m/z 576 and m/z 158. The presence of desaminyl-erythronolide A (m/z 576) indicates that the difference in m/z of 30 compared with erythromycin A (m/z 734) or of 16 compared with erythromycin C (m/z 720) is carried by the neutral sugar residue. The proposed structure for M1 is 3"-C desmethyl-2",3"-ene-erythromycin C.
- M2 produces a parent peak at m/z 706 and fragmentation products at m/z 576 and m/z 158. The presence of desaminyl-erythronolide A (m/z 576) indicates that the difference in m/z of 28 compared to erythromycin A (m/z 734) or of 14 compared to erythromycin C (m/z 720) is carried by the neutral sugar residue. The proposed structure for M2 is 3"-C desmethyl-erythromycin C.
- M3 produces a parent peak at to m/z 690 and fragmentation products at m/z 560 and m/z 158. The presence of desaminyl-erythronolide B (m/z 560) indicates that the difference in m/z of 28 compared to erythromycin B (m/z 718) or of 14 compared to erythromycin D (m/z 704) is carried by the neutral sugar residue. The proposed structure for M3 is 3"-C desmethyl-erythromycin D.
- M4 produces a parent peak at m/z 720 and fragmentation products at m/z 576 and m/z 158. The profile is identical to that of erythromycin C (m/z 720) with the presence of desosaminylerythronolide A (m/z 576) and the loss of the aminated sugar residue (m/z 158), but the metabolite M4 does not have the same retention time in RP-HPLC as erythromycin. The proposed structure for M4 is 3"-C desmethyl-erythromycin A.

The detection by MS-MS of the minor metabolite M1 having

an unsaturated neutral sugar (3"-C desmethyl-2",3"-ene-erythromycin C) indicates that the *eryBII* gene codes for dTDP-4-keto-L-6-desoxyhexose 2,3-reductase in the biosynthesis route of dTDP-mycarose.

- 5 The BII92 strain was deposited at the Collection Nationale de Cultures de Microorganismes (CNCM) INSTITUT PASTEUR, 25, Rue du Docteur Roux 75724 PARIS CEDEX 15 FRANCE, on the 16 July 1997 under the number I-1903.

EXAMPLE 5: construction of the plasmid pCIIIA.

- 10 ORF8 may be translationally coupled to ORF7 situated downstream, an in phase deletion was introduced so as to avoid a polar effect. An integration plasmid, called pCIIIA which carries such a deletion, was constructed according to the diagram in Figure 7(A-D).
- 15 An *SalI* deletion of 663 bp was introduced into ORF8 from nucleotide 9384 to nucleotide 10046 of the sequence of Figure 2 by sub-cloning in the plasmid pUC19 the two *SalI* fragments (a: 794 bp and b: 631 bp shown in Figure 5A) isolated from the plasmid pBK44 obtained in Example 1 in order to generate
- 20 the plasmid pdel88 (Figure 7A). The presence of the deletion of 663 bp was confirmed by sequencing. The plasmid pdel88 was then subjected to two additional sub-clonings so as to extend the chromosomal regions which can be used for the homologous recombination of the two sides of the deletion
- 25 site. The *SacI* fragment (450 bp) of the plasmid pdel88 was first replaced by the *SacI* fragment (1.1 kb) of the plasmid pEco2 obtained in Example 1 in order to generate the plasmid pdel88A (Figure 7B). Then the *EcoRI* fragment (1.5 kb) carrying the deletion in ORF8 was isolated from the plasmid
- 30 pdel88A and used to replace the *EcoRI* fragment (1.66 kb) carrying the intact ORF in the plasmid pOBB. The plasmid pOBB, represented in Figure 7C, corresponds to the plasmid pBK44 prepared in Example 1 in the site *PstI* of which there was sub-cloned the fragment *PstI* of 4 kb of the plasmid
- 35 pIJ486 obtained by digestion by the *PstI* restriction enzyme and carrying of the streptomyces replication origin as well as the thiostrepton resistance gene. The resultant plasmid pCIIIA carries chromosomal regions for the homologous

002220 50429450

recombination of 1.27 kb and 1.38 kb upstream and downstream respectively of the deletion site. The plasmid pCIIIA thus obtained (Figure 7D) was then transferred into the *E. coli* DH5 α MRC strain, then used for transforming *Sac. erythraea*.

5 **EXAMPLE 6: construction of a *Sac. erythraea eryCIIIA* strain (CIII68).**

10 A strain in which the *eryCIII* gene carries an internal deletion such as that introduced into the plasmid pCIIIA obtained in Example 5 was prepared by transformation of the protoplasts of *Sac. erythraea* with the plasmid pCIIIA.

The preparation of the protoplasts, the integration process and the selection of the mutants having the *ery*⁻ phenotype were carried out as in Example 3.

15 In addition, the presence of the expected deletion in the chromosome (deletion of 663 bp from nucleotide 9384 to nucleotide 10046 of the sequence of Figure 2) was confirmed by genomic analysis by Southern blot as well as by PCR amplification according to the conditions described in Example 3.

20 Using Southern hybridization on the genomic DNA digested by the *EcoRI* restriction enzyme, using as probe the oligonucleotide having the following sequence
C3-S ATGCGCGTCGTCTTCTCCTCCATG (SEQ ID No.32)
corresponding to the complementary strand of the DNA region
25 situated from position 10196 to position 10219 of the sequence in Figure 2, a band of 2.2 kb from the wild strain and a band of 1.5 kb from the mutant CIII68 were detected. The results shown in Figure 15 indicate the presence in the mutant of a deletion of approximately 700 bp in this region
30 of the chromosome.

The PCR amplification was carried out using the above oligonucleotide C3-S and the oligonucleotide having the following sequence

35 C3-R TCATCGTGGTTCTCTCCTTCC (SEQ ID No.33)
corresponding to the sequence situated from position 8954 to position 8974 of the sequence in Figure 2 allowing framing by PCR amplification of the region carrying the internal deletion at ORF8. Analysis by PCR amplification allowed a

band of approximately 1.2 kb in the wild strain and a band of approximately 0.6 kb in the mutant CIII68 to be detected in an identical fashion to the signal obtained with pBIIA. The results shown in Figure 16 confirm that the deletion of approximately 700 bp detected by the Southern analysis is identical to that carried by the plasmid pCIIIA (633 bp).

The recombinant strain thus obtained, designated CIII68, was then cultivated in order to identify the metabolites produced by the strain.

10 **EXAMPLE 7: fermentation of the CIII68 strain and identification of the secondary metabolites produced.**

The culture of the CIII68 strain and the analyses by TLC followed by bioautography were carried out according to the conditions indicated in Example 4.

15 The TLC results (Figure 17) show that the CII68 strain preferentially accumulates 3- α -mycarosyl erythronolide B as well as small quantities of erythronolide B as expected of an eryC mutant.

The eryCIII sequence has a strong homology with other putative glycosyltransferases such as DauH (43 % identity at the protein level) and DnrS (47 % identity) involved in the biosynthesis of daunorubicin in *S. peucetius* (Otten et al., 1995) and in *Streptomyces* sp C5 (Dickens et al., 1996) as well as TylM2 (50 % identity) involved in the transfer of mycaminoase onto tylactone in the biosynthesis route of tylosine in *S. fradiae* (Gandecha et al., 1997).

20 These observations indicate that eryCIII gene codes for desosaminyltransferase in the biosynthesis route of erythromycin.

30 **EXAMPLE 8: construction of the plasmid pCIIA .**

An integration plasmid, called pCIIA and carrying a deletion in the eryBII gene coding for ORF9, was constructed according to the diagram in Figure 8A.

35 The plasmid pK23 (Figure 5A) was obtained by sub-cloning in pUC19 of the *Kpn*I fragment of 10 kb isolated from the DNA of the clone λ SE5.5 digested by the *Kpn*I restriction enzyme.

Firstly, the shuttle vector pORT1, shown in Figure 8B, was obtained by sub-cloning the *Pst*I fragment of 4kb isolated

An out-of-phase deletion of 304 bp was introduced into

EXAMPLE 9: construction of a *Sac. erythraea* eryCIIΔ strain (CII62).

The preparation of the protoplasts, the integration process and the selection of the mutants having the *ery^r* phenotype were carried out as in Example 3.

Using Southern hybridization on the genomic DNA digested by the *Eco*RI restriction enzyme, using as probe the oligonucleotide C3-S having the above sequence, a band of 2.2 kb from the wild strain and a band of 1.8 kb from the mutant CII62 were detected. The results shown in Figure 15 indicate the presence in the mutant of a deletion of approximately 400 bp in this region of the chromosome.

The PCR amplification was carried out using the
35 oligonucleotide having the following sequence
C2-S GGAATTCATGACCACGACCGATC (SEQ ID No.34)
corresponding to the complementary strand of the DNA region
of the end of the *eryAIII* gene situated from position 20258

to position 20280 of the sequence deposited in the EMBL base under the reference X62569 and described by Bevitt et al., 1992 and the oligonucleotide having the following sequence
 C2-R CGCTCCAGGTGCAATGCCGGGTGCAGGC (SEQ ID No.35)

5 corresponding to the sequence situated from position 10558 to position 10585 of the sequence of Figure 2 allowing framing by PCR amplification of the region carrying the internal deletion at ORF9. Analysis by PCR amplification allowed a band of approximately 760 bp in the wild strain and a band of
 10 approximately 460 bp in the mutant CII62 to be detected in an identical fashion to the signal obtained with the plasmid pCIIA. The results shown in Figure 16 confirm that the deletion of approximately 400 bp detected by the Southern analysis is identical to that carried by the plasmid pCIIA
 15 (304 bp).

The recombinant strain thus obtained, designated CII62, was then cultivated in order to identify the metabolites produced by the strain.

EXAMPLE 10: fermentation of the CII62 strain and identification of the secondary metabolites produced.

20 The culture of the CII62 strain and the analyses by TLC followed by bioautography were carried out according to the conditions indicated in Example 4.

The TLC results (Figure 17) show that the CII62 strain
 25 preferentially accumulates 3- α -mycarosyl erythronolide B as well as small quantities of erythronolide B as expected of an eryC mutant.

The eryCII sequence has a strong homology with genes involved in the biosynthesis routes for daunosamine (DnrQ, 38
 30 % identity at the protein level, Otten et al., 1995) and mycaminose (protein coded by ORF1*, 40 % identity at the protein level, Gandeche et al., 1997) which also need to transfer a keto group in position 3 from an adjacent carbon.

These observations indicate that the eryCII gene codes
 35 for dTDP-4-keto-D-6-desoxyhexose 3,4-isomerase in the biosynthesis route of dTDP-desosamine.

EXAMPLE 11: cloning and sequencing of the eryAI-eryK region of the cluster of biosynthesis genes of erythromycin.

13.2 kb comprising the totality of the *eryAI* gene and corresponding to the DNA region comprised between the *NcoI* site situated at position 44382 of the sequence of Figure 3 and the *NcoI* site situated at position 392 of the X62569 sequence (Bevitt et al., 1992). The probe was prepared in the following fashion: Firstly, the *NcoI* fragment of 13.2 kb was isolated from the plasmid pBK25 described by Bevitt et al., 1992 and sub-cloned in the *SmaI* site of pUC18 after filling in the *NcoI* ends with the Klenow fragment. From the plasmid pNC012 thus generated, the fragment of 13.2 kb was isolated by digestion with the *NcoI* restriction enzyme.

The plasmid pNCO28 was sequenced by the generation of sub-clones using exonuclease III according to the instructions given by the supplier for the Erase-a-Base Kit (Promega) by digesting with the restriction enzymes *SacI/XbaI* and *NsiI/BamHI* respectively for the reverse direction. The sequence was completed using as primers the synthetic oligonucleotides having the following sequences

For the sequencing of plasmid pNCO62, templates were generated by sonication of the DNA according to Bankier et al. (1987) using as primers pUC18 such as vector. The sequence was completed using the synthetic oligonucleotides having the following sequences:

646 CATCGTCAAGGAGTTTCGACGGT (SEQ ID No.39)
647 TGCGCAGGTCCATGTTACCCACGTT (SEQ ID No.40)
648 GCTACGCCCTGGAGAGCCTG (SEQ ID No.41)

649 GTCGCGGTCGGAGAGCACGAC (SEQ ID No.42) and
874 GCCAGCTCGGCGACGTCCATC (SEQ ID No.43).

The *Nco*I junctions were sequenced using as template the DNA of the cosmid cos6B obtained above the regions of which covering the *Nco*I sites were sequenced using the primers having the sequences 644 and 645 indicated above.

In addition, a **ClaI-NcoI** fragment of 0.9 kb, containing the start of the sequence of the *eryAI* gene and the 5' part of ORF13, was cloned in pUC18. This fragment was prepared in the following fashion: The plasmid pBK6-12 represented in Figure 5B was first generated by sub-cloning in the phagmide pTZ18R of the *KpnI* fragment of 4.5 kb isolated from the plasmid pBK25 described by Bevitt et al., 1992. The sub-clone pCN9 was then generated by sub-cloning of the *ClaI-NcoI* fragment of 0.9 kb isolated from the plasmid pBK6-12 in the *SmaI* site of pUC19, after filling in the ends using the Klenow fragment. The plasmid pCN9 thus obtained (Figure 5B) was sequenced. Templates were generated by sonication of the DNA according to Bankier et al. (1987) using pUC18 as vector. The sequence was completed using as probe the oligonucleotide having the following sequence:

875 CGACGAGGTCGTGCATCAG (SEQ ID No.44).

The sequencing of the DNA is carried out by the Sanger method (1977) using an automated sequencer on the double strand DNA templates with the Applied Biosystem 373 A sequencer. The assembling of the sequence data was carried out with SAP software (Staden, 1984). The sequences were analyzed using GCG software (Devereux, 1984).

The nucleotide sequences obtained allowed the nucleotide sequence of 8160 bp of Figure 3 (sequence of SEQ ID No.6) to be established in which seven ORFs (13-19) were identified respectively from nucleotide 43841 to nucleotide 44806, from nucleotide 44809 to nucleotide 46053, from nucleotide 46109 to nucleotide 46819, from nucleotide 46907 to nucleotide 48436, from nucleotide 48436 to nucleotide 49638, from nucleotide 49679 to nucleotide 51145 and from nucleotide 51177 to nucleotide 51755 (numbered in Figure 3 from the *Bam*HI site situated at the 5' end of the *ermE* gene)

(respectively sequence of SEQ ID No.6 from nucleotide 242 to nucleotide 1207, from nucleotide 1210 to nucleotide 2454, from nucleotide 2510 to nucleotide 3220, from nucleotide 3308 to nucleotide 4837, from nucleotide 4837 to nucleotide 6039, from nucleotide 6080 to nucleotide 7546 and from nucleotide 7578 to nucleotide 8156) and corresponding respectively to the *eryBIV*, *eryBV*, *eryCVI*, *eryBVI*, *eryCIV*, *eryCV* and *eryBVII* genes, according to Liu and Thorson (1994) the functional characterizations of which had not yet been identified. The seven ORFs (13-19) are in the same direction, reading being carried out from the 5' region of the *eryAI* gene.

Specimens of *E. coli* XL1-blue containing the coding region of the above ORFs were deposited at the Collection Nationale de Cultures de Microorganismes (CNCM) INSTITUT PASTEUR, 25, Rue du Docteur Roux 75724 PARIS CEDEX 15 FRANCE, on the 16 July 1997:

- the plasmid pBK6-12 comprising the coding sequence for ORF13 and for part of ORF14 under the number I-1898
- the plasmid pNCO28 comprising the coding sequence for ORFs 14 and 15 as well as for part of ORFs 13 and 16 under the number I-1901 and
- the plasmid pNCO62 comprising the coding sequence for ORFs 17, 18 and 19 as well as for part of ORF16 under the number I-1900.

25 **EXAMPLE 12: construction of the plasmid pBIVA.**

ORF13 being translationally coupled to ORF14 situated downstream, an in-phase deletion had to be introduced. An integration plasmid, called pBIVA and carrying this deletion, was constructed according to the diagram of Figure 9A.

30 The plasmid pPSP4 (Figure 5B) was first constructed by sub-cloning of the *PvuII*-*SpeI* fragment (2.7 kb) isolated from the plasmid pBK6-12 obtained in Example 11 and the *SpeI*-*PstI* fragment (1.6 kb) isolated from the plasmid pNCO28 obtained in Example 11 in the vector pUC19 digested beforehand using the *SmaI* and *PstI* restriction enzymes.

35 From the plasmid pPSP4, the plasmid p19BIVA was generated by deleting the *BclI*-*NcoI* fragment of 510 bp internal to ORF13 and by substituting for it 45 bp

originating from a synthetic adaptor of 54 bp. This adaptor was generated by pairing the 2 complementary oligonucleotides having the following sequences

SEQ A

5 AATTGATCAAGGTGAACACGGTCATGCGCAGGATCCTCGAGCGGAACTCCATGGGG
(SEQ ID No.45) and

SEQ B

CCCCATGGAGTTCCGCTCGAGGATCCTGCGCATGACCGTGTTCACCTTGATCAATT
(SEQ ID No.46)

10 creating a *Bcl*I site and an *Nco*I site framing the sequence of 45 bp.

For the pairing, the two oligonucleotides were brought to a 1.8 μ M final concentration in the hybridization buffer NaCl 50 mM, Tris, HCl 20 mM pH 7.4, MgCl₂.6H₂O 2 mM, heated
15 for 5 mn at 100°C then slowly cooled down to ambient temperature. After digestion with the *Nco*I and *Bcl*I restriction enzymes, a ligation was carried out in the plasmid pPSP4 the *Bcl*I-*Nco*I fragment of 510 bp of which had been eliminated beforehand. From the plasmid p19BIVA thus
20 generated, the *Sac*I-*Eco*RI fragment (2.2 kb) carrying the modified ORF13 was sub-cloned in the plasmid pUWL218 digested beforehand with the *Sac*I and *Eco*RI restriction enzymes. The integration plasmid pBIVA thus obtained (Figure 9B) was then transferred into the *E. coli* DH5 α MRC strain, then used to
25 transform *Sac. erythraea*.

EXAMPLE 13: construction of a *Sac. erythraea* eryBIVA strain (BIV87) .

A strain in which the *eryCIII* gene carries an internal deletion such as that introduced into the plasmid pBIVA
30 obtained in Example 12 was prepared by transformation of the protoplasts of *Sac. erythraea* with the plasmid pBIVA.

The preparation of the protoplasts, the integration process and the selection of the mutants having the *ery*⁻ phenotype were carried out as in Example 3.

35 In addition, the presence of the expected deletion in the chromosome (deletion of 510 bp from nucleotide 43872 to nucleotide 44382 of the sequence of Figure 3) and its replacement by the synthetic sequence of 45 bp was confirmed

Using Southern hybridization on the genomic DNA digested by the *Xho*I restriction enzyme, using as probe the B4-R oligonucleotide having the following sequence

B4-R AACTCGGTGGAGTCGATGTCGTCGCTGCGGAA (SEQ ID No.47)
corresponding to the complementary strand of the DNA region
situated from position 44687 to position 44718 of the
sequence of Figure 3, a band of 5.4 kb from the wild strain
and a band of 2.7 kb from the mutant BIV87 were detected.
The results shown in Figure 15 indicate the presence of an
additional XhoI site at a distance of 2.7 kb upstream of the
XhoI site situated at position 47114 of the sequence of
Figure 3 thus confirming the incorporation of the adaptor in
the mutant chromosome, as expected by the incorporation of
the above synthetic adaptor used to generate the plasmid
pBIVΔ.

The PCR amplification was carried out using the oligonucleotide having the following sequence

20 B4-S CAATATAGGAAGGATCAAGAGGTTGAC (SEQ ID No.48)

corresponding to the DNA region situated from position 43652 to position 43678 of the sequence of Figure 3 and the oligonucleotide B4-R having the sequence indicated above, allowing framing by PCR amplification of the region carrying

25 the internal deletion at ORF13. Analysis by PCR amplification allowed a band of approximately 1 kb in the wild strain and a band of approximately 500 bp in the mutant BIV87 to be detected in an identical fashion to the signal obtained with the plasmid p BIV87 (Figure 16).

30 The whole of the Southern analysis and PCR results confirm the presence of the deletion of 510 bp and of the synthetic adaptor at the level of the chromosome of the mutant.

The recombinant strain thus obtained, designated BIV87,
35 was then cultivated in order to identify the metabolites
produced by the strain.

EXAMPLE 14: fermentation of the BIV87 strain and identification of the secondary metabolites produced.

The culture of the BIV87 strain and the analyses by TLC were carried out according to the conditions indicated in Example 4.

The TLC results (Figure 17) show that the BIV87 strain preferentially accumulates erythronolide B as expected of an eryB mutant.

Minor metabolites with low mobilities showing an antibiotic activity were also detected. These metabolites were extracted and analyzed by RP-HPLC coupled with mass spectrometry as described in Example 4.

The mass spectrum results show that modified forms of erythromycin A, B, C and D were produced. A major metabolite and 3 minor metabolites were detected.

The major metabolite M5 produces a parent peak at m/z 702 with dehydration and fragmentation products at m/z 684, m/z 560 and m/z 158 and corresponds to the elimination of 2 hydrogen atoms in erythromycin D (m/z 704, m/z 686). The presence of desosaminy l erythronolide B (fragment m/z at 560) indicates that the difference in mass is carried by the neutral sugar. The proposed structure for this metabolite is 4"-keto erythromycin D.

The minor metabolites also produce a profile with a difference of 2 in the m/z values respectively:

- M6 (m/z at 718, m/z 700, m/z 576, m/z 158) instead of m/z 720, m/z 702 for erythromycin C;
- M7 (m/z at 732, m/z 714, m/z 576, m/z 158) instead of m/z 734, m/z 716 for erythromycin A;
- M8 (m/z at 716, m/z 698, m/z 560, m/z 158) instead of m/z 718, m/z 700 for erythromycin B.

The proposed structures are respectively 4"-keto erythromycin C for M6, 4"-keto erythromycin A for M7 and 4"-keto erythromycin B for M8.

These observations indicate that the eryBIV gene codes for dTDP-4-keto-L-6-desoxyhexose 4-reductase in the biosynthesis route of dTDP-mycarose.

The BIV87 strain was deposited at the Collection Nationale de Cultures de Microorganismes (CNCM) INSTITUT PASTEUR, 25, Rue du Docteur Roux 75724 PARIS CEDEX 15 FRANCE,

on the 16 July 1997 under the number I-1904

EXAMPLE 15: construction of the plasmid pBVΔ.

An integration plasmid, called pBVΔ and carrying a deletion in the eryBV gene coding for ORF14, was constructed according to the diagram in Figure 10A.

A deletion of 726 bp was generated in ORF14 from nucleotide 44963 to nucleotide 45688 of the sequence of Figure 3 by ligation of the *Bcl*I-KpnI fragment (1.1 kb) isolated from the plasmid pBK6-12, obtained in Example 11, with the *Kpn*I-BamHI fragment (1.1 kb) isolated from the plasmid pNCO28, obtained in Example 11, in the plasmid pUWL218 digested beforehand by the *Bam*HI restriction enzyme. The integration plasmid pBVΔ thus obtained (Figure 10B) was then transferred into the *E. coli* DH5αMRC strain, then used to transform *Sac. erythraea*.

EXAMPLE 16: construction of a *Sac. erythraea* eryBVΔ strain (BV88).

A strain in which the eryBV gene carries an internal deletion such as that introduced into the plasmid pBVΔ obtained in Example 15 was prepared by transformation of the protoplasts of *Sac. erythraea* with the plasmid pBVΔ.

The preparation of the protoplasts, the integration process and the selection of the mutants having the ery⁻ phenotype were carried out as in Example 3.

In addition, the presence of the expected deletion in the chromosome (deletion of 726 bp from nucleotide 44963 to nucleotide 45688 of the sequence of Figure 3) was confirmed by genomic analysis by Southern blot as well as by PCR amplification according to the conditions described in Example 3.

Using Southern hybridization on the genomic DNA digested by the *Nco*I restriction enzyme, using as probe the oligonucleotide having the following sequence:
B5-R TCCGGAGGTGTGCTGTCGGACGGACTTGTCGGTCGAAA (SEQ ID No.49)
corresponding to the complementary strand of the DNA region situated from position 46060 to position 46098 of the sequence of Figure 3, a band of 2.7 kb from the wild strain and a band of 2.0 kb from the mutant BV88 were detected. The

09463703 022000

results shown in Figure 15 indicate in the mutant the presence of a deletion of about 700 bp in this region of the chromosome.

The PCR amplification was carried out using the
 5 oligonucleotide having the following sequence:
 B5-S AGGAGCACTAGTGC GG TACTGCTGACGTCCTT (SEQ ID No.50)
 corresponding to the DNA region situated from position 44799
 to position 44831 of the sequence of Figure 3 and the
 oligonucleotide B5-R having the sequence indicated above,
 10 allowing framing by PCR amplification of the region carrying
 the internal deletion at ORF14. Analysis by PCR amplification
 has allowed a band of approximately 1.3 kb in the wild strain
 and a band of approximately 570 bp in the mutant BV88 to be
 detected in an identical fashion to the signal obtained with
 15 the plasmid pBV88 (Figure 16). These results confirm that
 the deletion of 710 bp detected by Southern analysis is
 identical to that carried by the plasmid pBVA (726 bp).

The recombinant strain thus obtained, designated BV88,
 was then cultivated in order to identify the metabolites
 20 produced by the strain.

**EXAMPLE 17: fermentation of the BV88 strain and
 identification of the secondary metabolites produced.**

The culture of the BV88 strain and the analyses by
 TLC were carried out according to the conditions indicated in
 25 Example 4.

The TLC results (Figure 17) show that the BV88 strain
 preferentially accumulates erythronolide B as expected of an
 eryB mutant.

Minor metabolites with low mobilities were also detected
 30 then extracted and identified by RP-HPLC coupled with mass
 spectrometry under the conditions used in Example 4.

The mass spectrum shows the presence of a metabolite
 having a parent peak at m/z 560 and dehydration and
 fragmentation products at m/z 542 and m/z 158 the proposed
 35 structure for which is desosaminy l erythronolide B.

The eryBV sequence shows a strong homology with other
 glycosyltransferases as well as with the eryCIII gene above
 (60.7 % identity at the nucleotide level, 44 % at the protein

level).

These observations indicate that the *eryBV* gene codes for the mycarosyltransferase involved in the biosynthesis of erythromycin.

5 **EXAMPLE 18: construction of a plasmid pCVIA (pPSTI).**

An integration plasmid, called pPSTI and carrying a deletion in the *eryCVI* gene coding for ORF15, was constructed according to the diagram in Figure 11A in the following fashion:

10 Firstly, the plasmid pNB49 was generated by treatment with the exonuclease III of the plasmid pNCO28 obtained in Example 11 digested beforehand by the *NsiI* and *BamHI* restriction enzymes. The plasmid pNB49 (Figure 5B) containing the nucleotides 44382 to 46562 of the sequence of
15 Figure 3, was then digested with the help of the *PstI* restriction enzyme then treated with Mung Bean nuclease (NE Biolabs) as described by Sambrook et al. (1989). After religation and transformation in *E. coli* XL1-Blue, the ampicillin-resistant colonies were selected by restriction
20 analysis with the *PstI* enzyme. The loss of the *PstI* site was confirmed by sequencing a clone using the reverse M13 primer and deletion from nucleotide 46364 of the sequence of Figure 3 was observed creating a phase change in ORF15 in the plasmid pNB49Δ*Pst* thus generated. The plasmid pIJ702
25 digested with the *BglIII* restriction enzyme was then ligated with the *BglIII* site of the plasmid pNB49Δ*Pst* generating the plasmid pPSTI. The orientation of pIJ702 in pPSTI was confirmed by the presence of a DNA fragment having 0.9 kb after digestion with the *SphI* restriction enzyme. The
30 integration plasmid pPSTI (Figure 11B) thus obtained was transferred into the *E. coli* DH5αMRC strain, then used to transform *Sac. erythraea*.

EXAMPLE 19: construction of a *Sac. erythraea eryCVIA* strain (Pst10).

35 A strain in which the *eryBV* gene carries an internal deletion such as that introduced into the plasmid pPSTI obtained in Example 18 was prepared by transformation of the protoplasts of *Sac. erythraea* with the plasmid pPSTI.

002220 5022450

*Bgl*III site downstream of the sequence corresponding to the complementary strand of the DNA region situated from position 46027 to position 46053 of the sequence of Figure 3. The chromosomal DNA digested beforehand with the *Cla*I and *Pst*I restriction enzymes showed the expected bands of 4 kb and 7 kb from the integrant while the wild strain presented the expected band of 3 kb.

After repeated cultures of the integrants, the individualized colonies obtained were analyzed for sensitivity to thiostrepton and production of erythromycin, then the integration of the expected deletion (deletion from nucleotide 46364 of the sequence of Figure 3) into the chromosome of a mutant clone having the *ery*⁻ phenotype (*Pst*10) was confirmed by Southern analysis. The chromosomal DNA, isolated from the wild strain and the *Pst*10 mutant respectively, was digested with the *Pst*I restriction enzyme. Hybridization with the *Pst*I-*Nco*I probe of 0.8kb (nucleotides 46368 to 47142 of the sequence of Figure 3) produced the expected profile with a *Pst*I band of 1kb corresponding to nucleotides 46368 to 47397 of the sequence of Figure 3 from the wild strain and with a band of >20 kb from the mutant. The loss of the *Pst*I site at position 46368 above was also shown after double digestion by the *Pst*I and *Nco*I enzymes, resulting in a *Pst*I-*Nco*I band of 0.8 kb (nucleotide 46368 to 47142) from the wild strain and a *Nco*I band of 2.8 kb (nucleotide 44382 to 47142) with the mutant.

The recombinant strain thus obtained, designated *Pst*10, was then cultivated in order to identify the metabolites produced.

30 **EXAMPLE 20: fermentation of the *Pst*10 strain and identification of the secondary metabolites produced.**

The *Pst*10 strain was cultivated in the sucrose-succinate medium described by Caffrey et al. (1992) for 3 days at 30°C.

The culture supernatant was then extracted at pH 9 with ethyl acetate. The organic phases obtained were dried over *Mg*₂*SO*₄ then brought to dryness under reduced pressure. The residue was dissolved in the acetonitrile-water mixture (1:1, v/v), then analyzed by mass spectrometry on a BioQ

The production of erythromycin A (m/z 734 and m/z 716) was not observed but the presence of erythronolide B (MK+: m/z 441 and MNa+: m/z 425) as well as of 3- α -mycarosyl erythronolide B (MK+: m/z 585 and MNa+: m/z 569) demonstrated characterizes the Pst10 strain as an eryC mutant.

These observations indicate that the eryCVI gene codes for the dTDP-D-6-desoxyhexose 3-N-methyltransferase involved in the biosynthesis route of dTDP-D-desosamine.

An integration plasmid, called pXhoI and carrying a deletion in the *eryBVI* gene coding for ORF16, was constructed according to the diagram in Figure 12A in the following fashion:

25 The *Nco*I-*Xho*I fragment (3.1 kb) of the plasmid pNC062
obtained in Example 11 and containing the nucleotides 47142
to 50254 of the sequence of Figure 3 was sub-cloned in the
*Nco*I and *Xho*I sites of the plasmid Litmus 28. The plasmid
pNC062X (Figure 5B) thus generated was digested with the *Pst*I
30 restriction enzyme then treated with DNA polymerase T4
(Boehringer Mannheim). After religation and transformation
in *E. coli* XL1-Blue, the loss of the *Pst*I site at nucleotide
47397 of the sequence of Figure 3 was confirmed by sequencing
and a deletion of 60 bp from nucleotide 47337 to nucleotide
35 47397 of the sequence of Figure 3 were observed. The plasmid
pIJ702 digested with the *Bgl*II restriction enzyme was then
ligated to the *Bgl*II site of this construction. The
orientation of pIJ702 in the construction was confirmed by

5 EXAMPLE 22: construction of a *Sac. erythraea* strain eryBVIΔ
(Xho91).

The preparation of the protoplasts and the integration process were carried out as in Example 3.

The integration into the chromosome and the presence of the expected deletion were confirmed by Southern analysis according to the general methods described in Example 19.

After repeated cultures of the integrants, the individualized colonies obtained were analyzed for sensitivity to thiostrepton and production of erythromycin, then the integration of the expected deletion (deletion of 60 bp from nucleotide 47338 to nucleotide 47397 of the sequence of Figure 3) into the chromosome of a mutant clone having the *ery⁻* phenotype (*Xho*I) was confirmed by Southern analysis. The chromosomal DNA, isolated from the wild strain and the mutant *Xho*I respectively, was digested with the *Pst*I restriction enzyme. Hybridization with the *Pst*I-*Nco*I probe of 0.8 kb (nucleotides 46368 to 47142 of the sequence of Figure 3) produced the expected profile with a *Pst*I band of 1 kb corresponding to nucleotides 46368 to 47397 of the sequence

of Figure 3 from the wild strain and with a band of 4 kb from the mutant indicating that the *Pst*I site at position 47397 above was lost.

The loss of the *Pst*I site at position 47397 was also confirmed by PCR. The chromosomal DNA was subjected to an amplification by PCR using the primers corresponding respectively to the sequence from nucleotide 47300 to nucleotide 57320 and to the sequence from nucleotide 47661 to nucleotide 47636 of the sequence of Figure 3. An expected fragment of 306 bp was thus amplified from the wild strain generating after digestion with the *Pst*I restriction enzyme two bands of approximately 100 and 300 bp. From the mutant Xho91, a fragment of 300 bp was amplified, resulting from the deletion of 60 bp. This fragment was then isolated and was found to be resistant to digestion by the *Pst*I enzyme.

The recombinant strain thus obtained and designated Xho91, was then cultured in order to identify the metabolites produced.

EXAMPLE 23: fermentation of the Xho91 strain and identification of the secondary metabolites produced.

The culture of the Xho91 strain and the analysis of the culture supernatant by mass spectrometry were carried out according to the conditions described in Example 20.

The production of erythromycin A (m/z 734 and m/z 716) was not observed but the presence of a majority quantity of erythronolide B (MK^+ : m/z 441; MNa^+ : m/z 425; $M-H_2O$ H^+ : m/z 385) as well as the presence of desosaminyl erythronolide B (m/z 560) demonstrated characterize the Pst10 strain as an eryB mutant.

The mass spectrometry results were confirmed by high resolution mass spectrometry on a Bruker FT-ICR spectrometer (Bruker, FRG).

The eryBVI sequence has a strong homology with DnmT involved in the biosynthesis of daunorubicin in *S. peucetius* (accession number EMBL U77891) (43.9 % identity at the protein level).

These observations indicate that the eryBVI gene codes for the dTDP-4-keto-L-6-desoxyhexose 2,3-dehydratase

involved in the biosynthesis of dTDP-mycarose, as suggested by Scotti and Hutchinson, 1996.

EXAMPLE 24: construction of the plasmid pCIVA.

An integration plasmid, called pCIVA and carrying a deletion in the *eryCIV* gene coding for ORF17, was constructed according to the diagram in Figure 13A in the following fashion:

The plasmid pNCO62 obtained in Example 11 was digested using the *BalI* and *BclI* restriction enzymes so as to eliminate a fragment having 949 bp inside ORF17 from nucleotide 48650 to nucleotide 49598 of the sequence of Figure 3. After filling in the ends using the Klenow fragment of DNA polymerase I, the plasmid was religated and transformed in *E. coli* XL1-blue. From the plasmid pBCB17 thus generated, the fragment of 2.68 kb carrying the deletion was isolated by digestion using the *XbaI* and *SphI* enzymes, then sub-cloned in the sites corresponding to the plasmid pUWL218. The presence of the deletion of 949 bp from nucleotide 48650 to nucleotide 49598 of the sequence of Figure 3 was confirmed by sequencing. The integration plasmid pCIVA thus obtained (Figure 13B) was then transferred into the *E. coli* DH5 α MRC strain, then used to transform *Sac. erythraea*.

EXAMPLE 25: construction of a *Sac. erythraea* strain *eryCIVA* (CIV89).

A strain in which the *eryCIV* gene carries an internal deletion such as that introduced into the plasmid pCIVA obtained in Example 24 was prepared by transformation of the protoplasts of *Sac. erythraea* with the plasmid pCIVA.

The preparation of the protoplasts, the integration process and the selection of the mutants having the *ery*⁻ phenotype were carried out as in Example 3.

In addition, the presence of the expected deletion in the chromosome (deletion of 949 bp from nucleotide 48650 to nucleotide 49598 of the sequence of Figure 3) was confirmed by genomic analysis by Southern blot as well as by PCR according to the conditions described in Example 3.

Using Southern hybridization on the genomic DNA digested

05463705 022300

by the NcoI restriction enzyme, using as probe the oligonucleotide having the following sequence

C4-R AGCGGCTTGATCGTGTGGACCACTAC (SEQ ID No.53)

corresponding to the complementary strand of the DNA region situated from position 49996 to position 50022 of the sequence of Figure 3, a band of 6.2 kb from the wild strain and a band of 5.2 kb from the mutant CIV89 were detected. The results shown in Figure 15 indicate the presence in the mutant of a deletion of approximately 1 kb in this region of the chromosome.

The PCR amplification was carried out using the oligonucleotide having the following sequence:

C4-S GGCCTATGTGGACTACGTGTTGAACGT (SEQ ID No.54)

corresponding to the DNA region situated from position 48169 to position 48195 of the sequence of Figure 3 and the C4-R oligonucleotide having the sequence indicated above, allowing framing by PCR amplification of the region carrying the internal deletion at ORF17. Analysis by PCR amplification allowed a band of 1.8 kb in the wild strain and a band of 900 bp in the mutant CIV89 to be detected in an identical fashion to the signal obtained with the plasmid pCIVΔ. The results shown in Figure 16 confirm that the deletion of approximately 900 bp detected by Southern analysis is identical to that carried by the plasmid pCIVΔ (949 bp).

The recombinant strain thus obtained, designated CIV89, was then cultivated in order to identify the metabolites produced by the strain.

EXAMPLE 26: fermentation of the CIV89 strain and identification of the secondary metabolites produced.

The culture of the CIV89 strain and the analyses by TLC were carried out according to the conditions indicated in Example 4.

The TLC results (Figure 17) show that the CIV89 strain preferentially accumulates 3-α-mycarosyl erythronolide B as well as erythronolide B as expected of an eryC mutant.

Minor metabolites with low mobilities were also detected, then extracted and analyzed by RP-HPLC coupled with mass spectrometry according to the conditions used in Example

002220 5022450

4.

A minor metabolite produces a parent peak at m/z 720 and dehydration and fragmentation products at m/z 702, m/z 576 and m/z 174. Peak 174 may correspond to 4-hydroxydesosamine and peak 576 to 4'-hydroxydesosaminyl erythronolide B.

These results suggest that the difference in m/z of 16 compared with erythromycin D (parent peak m/z 704) is carried by the aminated sugar. The proposed structure for this metabolite is 4'-hydroxy erythromycin D.

These observations indicate that the enzyme is involved in the withdrawal of the hydroxyl group in the biosynthesis route of erythromycin and that the *eryCIV* gene codes for dTDP-6-desoxyhexose 3,4-deshydratase.

The CIV89 strain was deposited at the Collection Nationale de Cultures de Microorganismes (CNCM) INSTITUT PASTEUR, 25, Rue du Docteur Roux 75724 PARIS CEDEX 15 FRANCE, on the 16 July 1997 under the number I-1905.

EXAMPLE 27: construction of the plasmid pCVA .

An integration plasmid, called pCVA and carrying a deletion in the *eryCV* gene coding for ORF18, was constructed according to the diagram in Figure 14A in the following fashion:

The *BalI*-*BamHI* fragment (3.48 kb), obtained from the plasmid pNCO62 prepared in Example 11 by digestion with the *BalI* and *BamHI* restriction enzymes, was sub-cloned in the *SmaI*-*BamHI* sites of the vector pUC19. From the resulting plasmid pBAB18 (Figure 5B), the *ScaI* internal fragment (1kb) was then deleted by digestion with the *ScaI* enzyme in order to generate a deletion of 1044 bp from nucleotide 49998 to nucleotide 51041 of the sequence of Figure 3 in ORF18. From the plasmid pBABACV thus obtained, the fragment carrying the deletion was then reisolated from the polylinker of pUC19 by digestion with the *HindIII* and *EcoRI* restriction enzymes, then sub-cloned in the plasmid pUWL218. The integration plasmid pCVA thus obtained (Figure 14B) was transferred into the *E. coli* DH5 α MRC strain, then used to transform *Sac. erythraea*.

EXAMPLE 28: construction of a *Sac. erythraea* strain *eryCVA*

(CV90) .

A strain in which the *eryCV* gene carries an internal deletion such as that introduced into the plasmid pCVA obtained in Example 27 was prepared by transformation of the protoplasts of *Sac. erythraea* with the plasmid pCVA.

The preparation of the protoplasts, the integration process and the selection of the mutants having the *ery*⁻ phenotype were carried out as in Example 3.

In addition, the presence of the expected deletion in the chromosome (deletion of 1044 bp from nucleotide 49998 to nucleotide 51041) was confirmed by genomic analysis by Southern blot as well as by PCR according to the conditions described in Example 3.

Using Southern hybridization on the genomic DNA digested by the *Nco*I restriction enzyme, using as probe the oligonucleotide having the following sequence:
 C5-R AACGCCTCGTCCTGCAGCGGAGACACGAACA (SEQ ID No.55)
 corresponding to the complementary strand of the DNA region DNA situated from position 51229 to position 51259 of the sequence of Figure 3, a band of 6.2 kb from the wild strain and a band of 5.1 kb from the mutant CV90 were detected. The results shown in Figure 15 indicate in the mutant the presence of a deletion of approximately 1.1 kb in this region of the chromosome.

The PCR amplification was carried out using the oligonucleotide having the following sequence:
 C5-S TTCGCTCCCCGATGAACACAACCTCGTA (SEQ ID No.56)
 corresponding to the DNA region situated from position 49668 to position 49694 of the sequence of Figure 3 and the C5-R oligonucleotide having the sequence indicated above, allowing framing by PCR amplification of the region carrying the internal deletion at ORF18. Analysis by PCR amplification allowed a band of approximately 1.6 kb in the wild strain and a band of approximately 500 bp in the mutant CV90 to be detected in an identical fashion to the signal obtained with the plasmid pCVA. The results shown in Figure 16 confirm that the deletion of approximately 1.1 kb detected by the Southern analysis is identical to that carried by the plasmid

002220 6023450

PCVA (1044 bp).

The recombinant strain thus obtained, designated CV90, was then cultured in order to identify the metabolites produced by the strain.

5 **EXAMPLE 29: fermentation of the CV90 strain and identification of the secondary metabolites produced.**

The culture of the strain CV90 and the analyses by TLC were carried out according to the conditions indicated in Example 4.

10 The TLC results (Figure 17) show that the CV90 strain preferentially accumulates 3- α -mycarosyl erythronolide B as well as erythronolide B as expected of an *eryC* mutant.

The sequence of residues 38-50 (VTGAGDGDADVQA)
Val Thr Gly Ala Gly Asp Gly Asp Ala Asp Val Gln Ala
15 (SEQ ID No.61)
of the protein coded by *eryCV* (sequence of SEQ ID No.11) is close to the consensus sequence binding to NAD⁺ described by Wierenga et al., 1985 and by Scrutton et al., 1990.

20 These observations indicate that the *eryCV* gene codes for a reductase which would participate as a dTDP-4,6-desoxyhexose 3,4-reductase in the biosynthesis route of d-TDP-desosamine.

EXAMPLE 30: overexpression of the product of the *eryCIII* gene in *E. coli*.

25 The heterologous expression of the product of the *eryCIII* gene of *Sac. erythraea* corresponding to ORF8 described in Example 1 and coding for the desosaminyltransferase activity identified in Example 7 was carried out using *E. coli* as host strain. The protein thus
30 produced in the form of an inclusion body was then purified and its enzymatic activity determined in vitro.

1) Expression of the *EryCIII* protein in *E. coli*

The expression was carried out using the vector pET11a (Stratagène) for the cloning and the expression of
35 recombinant proteins in *E. coli* under the control of the promoter of the RNA polymerase of bacteriophage T7.

Firstly, the *eryCIII* gene was amplified from the plasmid pK62 described in Example 1 in the following fashion:

006220 50/0460

The PCR amplification was carried out using Native Pfu polymerase (Stratagène) and as primers the A oligonucleotide homologous with the coding strand of the *eryCIII* gene having the sequence

5 A GAAGGAGATATACATATGCGCGTCGTCTTCTCCTC (SEQ ID No.57)
allowing the introduction of an *NdeI* site upstream of the initiator ATG of *eryCIII* and the B oligonucleotide homologous with the complementary strand of the *eryCIII* gene having the sequence

10 B CGGGATCCTCATCGTGGTTCTCTCCTTCCTGC (SEQ ID No.58)
allowing the introduction of a *BamHI* site downstream of the stop codon of the *eryCIII* gene.

The amplified DNA was then digested by the *NdeI* and *BamHI* restriction enzymes, then the *NdeI*-*BamHI* fragment of
15 1.2 kb obtained containing the totality of the *eryCIII* gene was ligated in the expression vector pET11a (Stratagène) which contains the ampicillin-resistant β -lactamase gene, the replication origin *ColE1* and the promoter of the gene of RNA polymerase T7 situated upstream of the cloning site *NdeI*,
20 digested beforehand with the *NdeI* and *BamHI* restriction enzymes. After ligation and transformation in *E. coli* XL1-blue, the plasmid pCEIII thus obtained was confirmed by restriction map and sequencing.

The *E. coli* BL21 strain (DE3) of the pET kit
25 (Stratagène) which contains in its chromosomal DNA the *lacI^q* gene and the *lacUV5* promoter upstream of the gene of RNA polymerase T7, was then transformed by the plasmid pCEIII.5.

The transformed strain obtained, called BL21/pCEIII, was cultured in a 50 ml erlenmeyer flask at 37°C in LB medium
30 seeded at OD₆₀₀=0.1 from a preculture, then induced by isopropyl- β -D-thiogalactopyranoside (IPTG) 1mM at OD₆₀₀=1. After 3 hours 30 minutes of induction, 1 ml of culture was removed and centrifuged, then the bacterial pellet obtained was dissolved in 240 μ l of water and 120 μ l of SDS 3X sample
35 buffer (Tris-HCl 1M pH = 6.8: 1.9 ml; glycerol 3 ml; β -mercaptoethanol 1.5 ml; SDS 20 % , 3 ml; bromophenol blue 1 % pH = 7: 0.3 ml; H₂O sqf 10 ml). From 15 μ l of the solution obtained, the total proteins extracted were analyzed by SDS-

The overexpression of a protein having an apparent molecular weight of approximately 46 Kd corresponding to MW expected for the EryCIII protein was observed compared with the total proteins of a control strain transformed by the plasmid pET11a.

The transformed BL21/pECIII strain above was cultured in a 6 litre fermenter in minimum medium containing glycerol as the carbon source (Korz et al., 1995) at 25°C until OD₆₀₀=12, then induced by IPTG for 18 hours until OD₆₀₀=54. From the collected broth, the bacterial pellet containing inclusion bodies was isolated by centrifugation at 5000 g for 30 mn.

190 g of bacterial pellet corresponding to 1 litre of collected broth were resuspended in 2.5 volumes of $\text{KH}_2\text{PO}_4/\text{K}_2\text{HPO}_4$ 20 mM pH 7.2 buffer containing EDTA 2.5 mM and DTT 2.5 mM. The cells were then lysed using a Rannie apparatus (Mini-Lab, type 8-30H, APV Homogenisers As, Denmark) with three passes under a pressure of 1000 bar. After centrifugation at 46.000 g for 3 hours, the pellet obtained was suspended in 2.5 volumes of 2M urea then centrifuged under the same conditions.

The extract in the 7M urea was then loaded at a rate of 0.5 metres/h and at 4-8°C on a 180 ml column (5 cm x 9 cm) of

Q sepharose (Pharmacia) equilibrated beforehand with the buffer A above and with detection at 280 nm. The EryCIII protein was then eluted with buffer A containing 0.3M NaCl. The combined fractions, containing the EryCIII protein, demonstrated by SDS-PAGE (polyacrylamide gradient: 10 to 20 %) revealed by staining with Comassie blue and 835 mg of total proteins were then loaded onto a 5.5 litre column (10 cm x 70 cm) of Superdex 200 Prep grade (Pharmacia) equilibrated beforehand with buffer A above. By elution of the column with buffer A and by detection at 280 nm, a protein peak was obtained the fractions of which containing the EryCIII protein demonstrated by SDS-PAGE and 200 mg of total proteins were combined then purified on a 180 ml column (5 cm x 9 cm) of Q Source (Pharmacia) equilibrated beforehand with buffer A. By elution with a linear gradient of NaCl varying from 0 to 0.3 M in buffer A, 30 ml of solution containing 100 mg of denatured EryCIII protein of homogeneous purity evaluated by SDS-PAGE, with revelation with silver nitrate, were obtained.

Figure 18 shows the pattern of the purity of the EryCIII protein monitored by SDS-PAGE (polyacrylamide gradient: 10 to 15 %) for a deposit of 500 ng of total proteins and revelation with silver nitrate successively after extraction with 7M urea (line 2), Q sepharose chromatography (line 3), Superdex chromatography (line 4), Q source chromatography (line 6) relative to the molecular weight markers (lines 1 and 5).

The EryCIII protein was then renatured by dilution of the homogeneous eluate with a solution of buffer A containing DTT 10 mM in order to obtain a final protein concentration of 0.1 mg/ml. The diluted solution was then dialyzed against Tris buffer 50 mM; NaCl 0.15 M; 0.3 % n-octyl- β -D-glucopyranosyl (NOG); DTT 10 mM, pH 8.3 then concentrated to 4 mg/ml by ultrafiltration on a PLGC04310 membrane (Millipore) having a cut-off threshold of 10.000.

The purified EryCIII protein was then stored in a frozen state at -20°C in aliquots of 500 μl .

3) Characterization of the EryCIII protein

The characterization of the EryCIII protein thus obtained was examined for the following properties:

a) Homogeneity.

Electrophoresis by SDS-PAGE (polyacrylamide gradient: 10 to 15 %) using the Phast System apparatus (Pharmacia) and revelation with silver nitrate shows a purity greater than 99 % for a deposit of 2000 ng.

b) Molecular weight by electrophoresis and mass spectrometry.

Using electrophoresis, an apparent MW of 46 kDa was determined in agreement with the calculated MW of 45929.

The analysis by RP-HPLC coupled with mass spectrometry (HPLC: ESI-SM) produces a mass of 45934 uma.

c) N-terminal amino acid sequence

The N-terminal sequence was determined by microsequencing on a Model A492 protein microsequencer coupled with an HPLC analyser of PTH-amino acids (Applied Biosystems).

No secondary sequence was revealed for the 10 first residues which is in agreement with the amino acid sequence described in Figure 2 (sequence of SEQ ID No.5).

d) Biological activity

The desosaminyl transferase activity of the EryCIII protein was determined in vitro by demonstration of the formation of erythromycin D from dTDP-D-desosamine, the preparation of which is described hereafter and of 3- α -mycarosyl erythronolide B (MEB) the preparation of which is described above in Materials and General Methods.

The reaction medium contains 150 nmoles of dTDP-D-desosamine, 137.4 nmoles of MEB and 1 mg of EryCIII protein using the following operating conditions:

4.78 ml of Tris buffer 50 mM pH 7.3 (buffer B); 20 μ l of dTDP-D-desosamine, triethylamine salt (150 nmoles) in solution in buffer B containing EDTA 1 mM and PEFABLOC O, 4 mM (Merck); 100 μ l of MEB (137.4 nmoles) in solution in buffer B and 1 mg of EryCIII protein corresponding to 250 μ l of an aliquot of frozen solution obtained above are successively introduced into a glass tube provided with a screw top.

After homogenisation by Vortex, the stoppered tube is placed for 5 hours in a thermostatically-controlled bath at 30°C, then the pH is adjusted to 9-10 with 32 % NaOH then the reaction mixture is extracted 3 times with 5 ml of ethyl acetate. The extract obtained, brought to dryness under reduced pressure, then taken up in 100 µl of methylene chloride, is then analyzed by TLC under the conditions indicated in Example 4 using a methylene chloride/methanol mixture (90: 10,v/v) as eluent.

A control test (t = 0) the incubation of which is stopped immediately by the addition of NaOH, is carried out under the same conditions.

The results obtained by chemical revelation show the appearance of a less mobile product having an R_f close to the expected erythromycin D and for which a weak antibiotic activity is detected by direct autobiogram of the plates on *B. pumilus*. No biological activity is observed for the control test (Figure 19).

These results confirm that the EryCIII protein produced in *E. coli* and purified above has the expected desoaminyl transferase activity and was correctly renatured.

EXAMPLE 31: use of the sequence of the eryCIII gene as probe to isolate the oleG1 and oleG2 genes coding for glycosyltransferases in *S. antibioticus*.

1) cloning of the oleG1 and oleG2 genes

The sequence of the eryCIII gene of *Sac. erythraea* corresponding to ORF8 described in Example 1 coding for the desosaminyltransferase activity was used to prepare a hybridization probe and allowed the isolation of the homologous genes in the *S. antibioticus* ATCC 11891 strain producing oleandomycin by Southern hybridization.

The whole of the eryCIII gene was amplified by PCR from 6 ng of the plasmid pK62 obtained in Example 1 following the operating conditions described in Example 3 using native pfu polymerase (Stratagène) and, as primers the oligonucleotide having the following sequence:

eryCIII-1 CGGGTACCATGCGCGTCGTCTTCTCCTCCATG (SEQ ID No.59) comprising a KpnI restriction site in its 5' region and the

3' part of which corresponds to the complementary strand of the DNA region situated from position 10196 to position 10219 of the sequence of Figure 2 and the *eryCIII2* oligonucleotide having the following sequence:

- 5 *eryCIII-2* CGGGTACCTCATCGTGGTTCTCTCCTTCC (SEQ ID No.60)
comprising a *KpnI* site in its 5' region and the 3' part of which corresponds to the DNA region situated from position 8954 to position 8974 of the sequence of Figure 2.

- The band of approximately 1.2 kb obtained by
10 amplification was then digested by the *KpnI* restriction enzyme and cloned in the plasmid pUC19 digested beforehand by the *KpnI* enzyme. The plasmid pCIIIPCR1 thus obtained was then used to reisolate the *KpnI* fragment of 1.2 kb corresponding to all of the *eryCIII* gene shown in Figure 2.
15 The fragment thus isolated was then labelled with ^{32}P using the "random priming" technique described by Sambrook et al., 1989 and used as an *eryCIII* probe to analyze by Southern hybridization cosmid clones obtained from a genomic DNA library of *S. antibioticus* ATCC 11891 and prepared in the
20 following fashion (Figure 20):

- A series of six cosmids (*cosAB35*, *cosAB76*, *cosAB87*, *cosAB67*, *cosAB63* and *cosAB61*) overlapping and covering approximately 100 kb of the region corresponding to the cluster of oleandomycin biosynthesis genes was isolated by
25 following the method described by Swan et al., 1994 and using as probes the *SmaI* internal fragment of 2 kb at the third sub-unit of the polyketide synthase of *Sac. erythraea* in the cluster of biosynthesis genes of erythromycin (Cortes et al., 1990) followed by chromosome walking. The *strD*, *strE* and
30 *strM* probes coding respectively for dTDP-glucose synthase, dTDP-glucose 4,6-deshydratase and dTDP-6-desoxyglucose 3,5-epimerase of *S. griseus* (Stockmann and Piepersberg, 1992) hybridize with the *cosAB61* and *cosAB63* cosmids (Fig. 20). In a similar fashion, by Southern hybridization with the *eryCIII*
35 probe prepared above carried out under the standard conditions described by Hopwood et al., 1985, the *cosAB35* cosmid (Swan et al., 1994) produces positive signals in two *BamHI* restriction fragments of 3.5kb and 2.7 kb represented

in Figure 20. Subsequent sub-cloning and sequencing shows that these two fragments are separated by a *Bam*HI fragment of 0.6 kb which is not detected by hybridization.

An SstI fragment of 10.8 kb of genomic DNA of *S. antibioticus* ATCC 11891 represented in Figure 21, corresponding to the right part of the cluster of oleandomycin biosynthesis genes comprised between the SstI site in position 11081 of the OLE-ORF3 gene of the PKSs of the EMBL No.L09654 sequence) and the SstI site in position 5 of the EMBL No.L36601 sequence situated 1.4 kb upstream of the oleB gene and hybridizing with the eryCIII probe prepared above, was starting isolated from the cosAB35 cosmid and sub-cloned in the plasmid vector pSL1180 (Pharmacia Biotech). The pCO35-S clone thus obtained was used to generate single-stranded templates by sub-cloning different DNA fragments in the M13mp18 and MP13mp19 bacteriophages (New England Biolabs), then the nucleotide sequence of these fragments was determined according to the method of Sanger et al. (1977) using a modified T7 polymerase (Sequenase version 2.0; U.S. Biochemicals) in the presence of α [³⁵S]dCTP (Amersham) and 7-deaza-dGTP, according to the supplier's recommendations in order to limit band compression problems. The standard primers supplied with the Sequenase kit as well as the internal synthetic primers(17mer) were used.

Assembly of the sequence data was carried out using the Fragment Assembly program (Genetic Computer Group, University of Wisconsin) and identification of the open reading frames using the CODONPREFERENCE program (Devereux et al., 1984).

The nucleotide sequences obtained allowed the nucleotide sequence of 6093 bp represented in Figure 22(sequence of SEQ ID No.15) to be established, comprised between the *Sph*I* and *Kpn*I sites shown in Figure 21, in which five ORFs were identified from nucleotide 184 to nucleotide 1386 (ORF called oleP1), from nucleotide 1437 to nucleotide 2714 (ORF called oleG1), from nucleotide 2722 to nucleotide 3999 (ORF called oleG2), from nucleotide 3992 to nucleotide 4729 (ORF called oleM) and from nucleotide 4810 to nucleotide 5967 (ORF called oleY) respectively). The five ORFs are transcribed in the

006220 5072460

Specimens of *E. coli* containing the plasmid pCO35-S comprising the coding region of ORFs *oleP1*, *oleG1*, *oleG2*, *oleM* and *oleY* were deposited at the Collection Nationale de Cultures de Microorganismes (CNCM) INSTITUT PASTEUR, 25, Rue du Docteur Roux 75724 PARIS CEDEX 15 FRANCE, on the 8 April 1998 under the number I-2003.

The *oleG1* gene codes for a polypeptide having 426 amino acids (sequence of SEQ ID No.17). However, the presence of a CGC codon coding for a highly preserved arginine in this class of glycosyltransferase in the *Streptomyces* situated immediately upstream of the GTG codon, would indicate that the initiation codon could be the CTG codon in position 1431 of the sequence SEQ ID No.17.

15 The *oleG2* gene codes for a polypeptide having 426 amino acids (sequence of SEQ ID No.18).

Comparison of the deduced amino acid sequences of the *oleG1* and *oleG2* ORFs above with the database proteins using the Blast program (Altschul et al., 1990) showed similarities with glycosyl transferases from different sources, in particular approximately 72 % similarity and 53 % identity with the EryCIII deosaminyltransferase described in Example 30.

Identification of the function of the *oleG1* gene or the *oleG2* gene was then carried out by interruption of the target gene in the *S. antibioticus* strain ATCC 11891 and by identification of a non-glycosylated precursor of the oleandomycin produced by the mutant strain using the methods described in Examples 3 to 4.

30 2) generation of an *S. antibioticus* oleG1Δ strain (A35G1).

A strain in which the *oleG1* gene is interrupted was prepared by integration of a plasmid pCO3 into the homologous regions of the chromosomal DNA of the of *S. antibioticus* strain ATCC 11891 producing oleandomycin.

35 Firstly, the *Bam*HI fragment of 0.6 kb internal to the *oleG1* gene, obtained by digestion of the plasmid pC035-S prepared above, with the *Bam*HI restriction enzyme (Figure 21), was sub-cloned in the *Bam*HI site of the plasmid pOJ260

NRRL B-14785.

The plasmid pCO3 thus generated was then transferred into the *E. coli* strain TG1 recO1504::Tn5 (Kolodner et al., 1985), then used to transform the protoplasts of *S.*

5 *antibioticus*. Selection of the transformants was carried out by resistance to apramycin (injection-grade Apramycin, Rhône Mérieux).

Preparation of the protoplasts was carried out from the *S. antibioticus* strain ATCC 11891 following the conditions
10 described by Hopwood et al., 1985.

Transformation was carried out using 50 µl of an aliquot of protoplasts, 5 µg of plasmidic DNA pCO3 and replacing the thiostrepton with apramycin at the final concentration of 25 µg/ml.

15 Selection of the integrants carried out by resistance to apramycin allowed a clone called A35G1 to be isolated.

The expected alteration in the corresponding region of the chromosome of *S. antibioticus* was confirmed by genomic analysis by Southern blot. The chromosomal DNA isolated,
20 then digested by the *Pst*I restriction enzyme from the wild *S. antibioticus* strain or the A35G1 mutant was analyzed by Southern using the *Bam*HI fragment of 0.6 kb indicated above as hybridization probe. Replacement of the *Pst*I fragment of 4.7 kb thus detected in the wild strain by two *Pst*I fragments
25 of 2.4 and 6.5 kb in the A35G1 mutant confirms the integration of the plasmid pCO3 into the chromosome of the A35G1 strain at the level of *ole*G1 ORF.

The A35G1 recombinant strain thus obtained was then cultured to identify the precursors produced by the strain.

30 **3) fermentation of the A35G1 strain and identification of the oleandomycin precursors produced.**

The A35G1 strain was cultured for 72 hours in a 50 ml erlenmeyer in EP2 medium from a 48 hour preculture in an EP1 medium under the conditions described in Example 4.

35 The broth extracts with ethyl acetate were then analyzed according to the methods used in Examples 3 and 4.

a) The biological test by record of bacterial sensitivity to antibiotics was carried out in the following

fashion:

After growth of the *B. pumilus* strain on TSB medium overnight at 37°C, the culture was diluted to 1/100 in medium containing 50 % (w/v) of glycerol, then the cell suspension
5 obtained was preserved at -20°C before use.

The biological test was then carried out by introducing 150 µl of the thawed cell suspension into 100 ml of TSB medium containing 1 % agar and kept at 55°C. The mixture was then poured into petri dishes. After cooling down, Oxford
10 cylinders containing 50 to 200 µl of ethyl acetate extracts were placed on agar dishes, kept for 2 hours at 4°C, then incubated overnight at 37°C.

The extracts show no inhibitory effect on the growth of *B. pumilus* ATCC 14884.

15 b) The analysis by TLC by chemical revelation was carried out according to the conditions described in Example 4 using as standards erythromycin A, erythronolide B as well as 6-desoxyerythronolide B.

The analysis by TLC shows that the A35G1 strain does not
20 produce oleandomycin but preferentially accumulates a crimson product having a mobility which is greater than erythronolide B and close to 6-desoxyerythronolide B and that the aglycone 8,8a-desoxyoleandolide part can be expected.

c) The analysis by RP-HPLC chromatography coupled with
25 mass spectrometry was carried out according to the conditions described in Example 4. Two major metabolites, called M9 and M10, were detected (elution at 6.12 mn and 17.23 mn respectively). The two products produce a parent peak at m/z 373 and similar fragmentation profiles which may be in
30 accordance with the structure [8,8a-desoxyoleandolide]H⁺. However, only the retention time of the M10 metabolite is in accordance with the proposed structure whilst the M9 metabolite could correspond to an isomeric structure or to the open lactone nucleus.

35 Experiments complementing the mutant strains of *Sac. erythraea* CIII68 described in Example 6 and BV88 described in Example 16 were also carried out using plasmidic constructions allowing respectively each of the *oleG1* and

006220 50/59460

oleG2 genes to be expressed.

These observations and the absence of detection of oleandrosyl 8,8a-desoxyoleandolide or desosaminyl 8,8a-desoxyoleandolide indicate that the oleG1 gene codes for the desoaminosyltransferase and the oleG2 gene codes for the oleandrosyltransferase respectively involved in the biosynthesis of oleandomycin.

PREPARATION OF EXAMPLE 30: Thymidine 5'-(trihydrogen-diphosphate), P'-[3,4,6-trideoxy-3-(dimethylamino)-D-xylohexopyranosyl]ester, N,N-diethylethanamine
STAGE A: 3,4,6-trideoxy-3-(dimethylamino)-D-xylohexopyranose hydrochloride

146.6 g of erythromycin A is added under agitation and at ambient temperature to 1.5 litres of 6N hydrochloric acid. The solution obtained is taken to reflux for 2 hours. The reaction medium is cooled down to ambient temperature, filtered and the residue obtained is washed with water. The aqueous phase is extracted with methylene chloride, then with sulphuric ether. 10 g of black L₂S is added to the aqueous phase and the traces of ether are driven off under reduced pressure. After filtering and concentrating, a second run is carried out under the same conditions. The two tests are combined, dissolved in ethanol (150 cm³), 150 cm³ of ethyl ether is added. The crystals obtained are separated, washed and dried. 42 g of sought product is obtained. M.p. = 158~160°C.

STAGE B: 3,4,6-trideoxy-3-(dimethylamino)-D-xylohexopyranose, 1,2-diacetate

60 cm³ of triethylamine is added under agitation at 20°C, to a mixture containing 15.27 g of the product of Stage A and 150 cm³ of methylene chloride. A solution containing 20 cm³ of acetic anhydride and 80 cm³ of methylene chloride is added at 20°C. Agitation is carried out at ambient temperature for 20 hours, followed by filtering, washing, concentrating and impasting in sulphuric ether. The filtrate is concentrated under reduced pressure. The residue obtained is chromatographed eluting with an ethyl acetate-triethylamine mixture (95-5). 18.6 g of sought product is obtained which

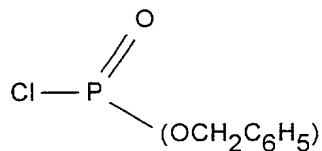
is used as such in the following stage.

STAGE C: 3,4,6-trideoxy-3-(dimethylamino)-D-xylo-hexopyranose,2-acetate

A mixture containing 18.6 g of the product of Stage B and 50 cm³ of DMF is taken to 50°C and 6.62 g of hydrazine acetate NH₂NH₂, ACOH is added. The reaction mixture is agitated and poured into a saturated solution of sodium acid carbonate. The aqueous phase is extracted with ethyl acetate. The organic phases are combined, dried, filtered and concentrated, followed by distillation under reduced pressure in order to eliminate the DMF by azeotropic entrainment with toluene. 11.28 g of product is obtained which is chromatographed on silica eluting with a ethyl acetate-triethylamine mixture (90-10). 6.5 g of sought product is obtained which is used as such in the following stage.

STAGE D: 3,4,6-trideoxy-3-(dimethylamino)-D-xylo-hexopyranose,2-acetate bis(phenylmethyl)phosphate

5.7 cm³ of a solution of n-butyllithium in hexane is added at -70°C to a solution containing 1.738 g of the product of the preceding stage and 40 cm³ of THF. 10 g of dibenzylphosphochloride prepared extemporaneously (J. Chem. Soc. 1958, p. 1957),



in solution in 20 cm³ of THF is added at -70~-75°C. Agitation is maintained for 1 hour 30 minutes between -70 and -74°C. The reaction medium is poured into a saturated solution of sodium acid carbonate and extracted with ethyl acetate. The organic phase is dried over sodium sulphate, filtered and concentrated. The product obtained is chromatographed on silica eluting with an acetone-methylene chloride mixture (5-5). 1.070 g of sought product is obtained.

STAGE E: 3,4,6-trideoxy-3-(dimethylamino)-D-xylo-hexopyranose,1-(dihydrogen phosphate),N,N-diethylethanamine

A mixture containing 1.070 g of the product of the preceding stage, 20 cm³ of ethyl acetate, 10 cm³ of methanol, 5 0.622 cm³ of triethylamine and 200 mg of palladium on carbon is placed under agitation and under a hydrogen flow for 30 minutes at ambient temperature. The reaction medium is filtered, washed with methanol and with ethyl acetate and the filtrate is concentrated. 1 g of an oil is obtained to which 10 10 cm³ of methanol is added. The solution obtained is agitated for 20 hours. The methanol is driven off under reduced pressure at 30°C. 680 g of sought product is obtained.

STAGE F: 3,4,6-trideoxy-3-(dimethylamino)-D-xylo- 15 hexopyranose,1-(dihydrogen phosphate)

420 mg of the product isolated in the form of triethylamine salts, obtained in the preceding stage is dissolved in 1.6 cm³ of methanol. 3.2 cm³ of sulphuric ether is added, then 6.4 cm³ of sulphuric ether is added. 250 mg of 20 sought product is obtained melting at 242~244°C.

STAGE G: Thymidine 5'-(trihydrogen diphosphate),P'-[3.4,6-trideoxy-3-(dimethylamino)-D-xylo-hexopyranosyl]ester,N,N-diethylethanamine

228 mg of the product of Preparation 1, 6 cm³ of pyridine 25 and 544 mg of thymidine 5'-monophosphate morpholidate-4-morpholine-NN'-dicyclohexylcarboxamide are mixed together. The pyridine is driven off under reduced pressure using a rotary evaporator while maintaining the temperature at 30°C or below. 6 cm³ of pyridine is added which is driven off 30 under reduced pressure. The operation is repeated twice. 6 cm³ of pyridine and 105 mg of 1H-tetrazole are added. Agitation is carried out for 3 days at ambient temperature. The pyridine is driven off under reduced pressure, followed by taking up in water, filtering, concentrating and a product 35 is obtained which is purified by chromatography. In this way the sought product is obtained.

rf = 0.12 eluent CH₂Cl₂, MeOH, H₂O (60-35-6).

002220 50/29460

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990)
J Mol Biol 215: 403-410.

- 5 Ausubel FM, Brent R, Kingston RE, Moore DD, Seidman JG, Smith JA Struhl K (1995) Current protocols in molecular biology. Massachusetts General Hospital and Harvard medical School. John Wiley and Sons, Inc.
- 10 Bankier AT, Weston KM, Barrell BG (1987) Methods in Enzymology 155: 51-93.
- Bevitt DJ, Cortes J, Haydock SF and Leadlay PF, (1992), Eur. J. Biochem 204: 39-49.
- 15 Caffrey P, Bevitt DJ, Staunton J, Leadlay PF (1992) FEBS 304: 225-228.
- 20 Cortés J, Haydock SF, Roberts GA, Bevitt DJ, Leadlay PF (1990) Nature 348: 176-178.
- Devereux J, Haeberli P, Smithies O (1984) Nucl Acids Res 12: 387-395.
- 25 Dhillon N, Hale RS, Cortes J, Leadlay PF (1989) Mol Microbiol 3: 1405-1414.
- Dickens ML, Ye J, Strohl WR (1996) J Bacteriol 178: 3384-3388.
- 30 Donadio S, Stassi J, McAlpine JB, Staver MJ, Sheldon PJ, Jackson M, Swanson SJ, Wendt-Pienkowski E, Wang YG, Jarvis B, Hutchinson CR, Katz L (1993) In: Baltz RH, Hegeman GD, Skatrud PL (eds) Industrial Microorganisms: Basic and Applied Molecular Genetics: American Society for Microbiology, Washington, DC, 257-265.

- Gandecha AR, Large SL, Cundliffe E (1997) Gene 184: 197-203.
- Haydock SF, Dowson JA, Dhillon N, Roberts GA, Cortes J, Leadlay PF (1991) Mol Gen Genet 230: 120-128.
- 5 Hessler PE, Larsen PE, Constantinou AI, Schram KH, Weber JM, Appl Microbiol. Biotechnol (1997), 47: 398-404.
- Hopwood DA, Kieser T, Wright HM and Bibb MJ, Journal of
10 General Microbiology (1983), 129, 2257-2269.
- Hopwood DA, Bibb MJ, Chater KF, Kieser T, Bruton C, Kieser HM, Lydiate DJ, Smith CP, Ward JM, Schremp H (1985) Genetic manipulation of Streptomyces. A laboratory Manual. John
15 Innes Foundation, Norwich.
- Kaneda T, Butte JC, Taubman B, Corcoran JW (1962) J Biol Chem 237: 322-327.
- 20 Katz E, Thompson CJ, Hopwood DA (1983) J Gen Microbiol 129: 2703-2714.
- Korz DJ, Rinas U, Hellmuth K, Sanders EA, Deckwer W-D (1995) Journal of Biotechnology 39: 59-65.
- 25 Katz L, Donadio S (1995) Macrolides. In: Vining LC, Stuttard C (eds). Genetics and Biochemistry of Antibiotic Production. Butterworth-Heinemann, Newton, MA, 385-420.
- 30 Liu H-W, Thorson JS (1994) Annu Rev Microbiol 48: 223-56.
- Otten SL, Liu X, Ferguson J, Hutchinson CR (1995) J Bacteriol 177: 6688-6692.
- 35 Sakakibara H and Omura S (1984). In: Omura S. (ed) Macrolide Antibiotics: Chemistry, Biology, and Practice. Academic Press, Inc. London, 85-125.

- Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, N.Y.
- 5 Sanger F, Nicklen S, Coulson AR (1977) Proc Natl Acad Sci USA 74: 5463-5467.
- Scotti C, Hutchinson CR (1996) J Bacteriol 178: 7316-7321.
- 10 Scrutton NS, Berry A, Perham RN (1990) Nature 343: 38-43.
- Staden R (1984), Nucleic Acids Res 12: 521-528.
- Stassi D, Donadio S, Staver MJ, Katz L (1993) J Bacteriol 175: 182-189.
- 15 Stemmer WPC (1994) Proc. Natl. Acad. Sci., Vol. 91, pp. 10747-10751.
- 20 Stockmann M, and Piepersberg W (1992) FEMS Microbiol Lett 90: 185-190.
- Swan D.G., Rodriguez A.M., Vilches C., Méndez C., Salas J.A., (1994) Mol Gen Genet 242: 358-362.
- 25 Ward JM, Janssen GR, Kieser T, Bibb MJ, Buttner MJ, Bibb MJ (1986) Mol. Gen. Genet. 203: 468-475
- Weber JM, Wierman CK, Hutchinson CR (1985) J Bacteriol 164: 425-433.
- 30 Weber JM, Losick R (1988) Gene 68: 173-180.
- Weber JM, Schoner B, Losick R (1989) Gene 75: 235-241.
- 35 Weber JM, Leung JO, Maine GT, Potenz RHB, Paulus TJ, DeWitt JP (1990) J Bacteriol 172: 2372-2383.

Weber JM, Leung JO, Swanson SJ, Idler KB, McAlpine JB (1991)
Science 252: 114-117.

Wehmeier UF (1995) Gene 165: 149-150.

5

Wierenga RK, Terpstra P, Hol WGJ (1986) J. Mol. Biol. 187:
101-107.

10 Yamamoto H, Maurer KH, Hutchinson CR (1986) J Antibiot 34:
1304-1313.

0946705-02100

Free text in the sequence listings

SEQ ID No.: 1:

/function= "involved in the biosynthesis of mycarose"

5 /gene= "eryBII"

/function= "involved in the biosynthesis of desosamine"

/gene= "eryCII"

SEQ ID No.: 4:

10 /function= "involved in the biosynthesis of desosamine"

/gene= "eryCIII"

/note= "SEQ ID No 1 FROM 1046 TO 2308"

SEQ ID No.: 6:

15 /function= "involved in the biosynthesis of mycarose"

/gene= "eryBIV"

/transl_except= (pos: 242 .. 244, aa: Met)

/function= "involved in the biosynthesis of mycarose"

/gene= "eryBV"

20 /transl_except= (pos: 1210 .. 1212, aa: Met)

/function= "involved in the biosynthesis of desosamine"

/gene= "eryCVI"

/function= "involved in the biosynthesis of mycarose"

/gene= "eryBVI"

25 /transl_except= (pos: 3308 .. 3310, aa: Met)

/function= "involved in the biosynthesis of desosamine"

/gene= "eryCV"

/function= "involved in the biosynthesis of mycarose"

/gene= "eryBVII"

30 /transl_except= (pos: 7578 .. 7580, aa: Met)

SEQ ID No.: 13:

/function= "involved in the biosynthesis of desosamine"

/gene= "eryCIV"

35 /note= "SEQ ID No 6 FROM 4837 TO 6039"

SEQ ID No.: 15:

/gene= "oleP1"

002220-5022950

```
/note= "SEQ ID No 11 FROM 38 TO 50"
```

CLAIMS

- 1) Isolated single- or double-stranded DNA sequence, represented in Figure 2 (direct or complementary sequence of SEQ ID No.1) corresponding to the eryG-eryAIII region of the cluster of erythromycin biosynthesis genes.
- 2) DNA sequence according to claim 1 comprising:
- the eryBII sequence corresponding to ORF7 (complementary sequence of SEQ ID No.1 from nucleotide 48 to nucleotide 1046) and coding for a dTDP-4-keto-L-6-desoxyhexose 2,3-reductase,
 - the eryCIII sequence corresponding to ORF8 (complementary sequence of SEQ ID No.1 from nucleotide 1046 to nucleotide 2308) and coding for a desosaminyltransferase and
 - the eryCII sequence corresponding to ORF9 (complementary sequence of SEQ ID No.1 from nucleotide 2322 to nucleotide 3404) and coding for a dTDP-4-keto-D-6-desoxyhexose 3,4-isomerase.
- 3) Isolated DNA sequence represented in Figure 2 chosen from the eryBII sequence corresponding to ORF7 (complementary sequence of SEQ ID No.1 from nucleotide 48 to nucleotide 1046), the eryCIII sequence corresponding to ORF8 (complementary sequence of SEQ ID No.1 from nucleotide 1046 to nucleotide 2308) or the eryCII sequence corresponding to ORF9 (complementary sequence of SEQ ID No.1 from nucleotide 2322 to nucleotide 3404) and the sequences which hybridize and/or display significant homologies with this sequence or fragments of the latter and having the same function.
- 4) Isolated DNA sequence eryCIII represented in Figure 2 corresponding to ORF8 (complementary sequence of SEQ ID No.1 from nucleotide 1046 to nucleotide 2308 = complementary sequence of SEQ ID No.4) and coding for a desosaminyl transferase.
- 5) Polypeptide coded by one of the DNA sequences according to one of claims 1 to 4.
- 6) Polypeptide according to claim 5 corresponding to an ORF represented in Figure 2, chosen from ORF7 (having the sequence of SEQ ID No.2), ORF8 (having the sequence of SEQ ID

No.5) or ORF9 (having the sequence of SEQ ID No.3) and the analogues of this polypeptide.

7) Polypeptide according to claim 5 corresponding to ORF8 represented in Figure 2 (having the sequence of SEQ ID No.5) and having a desosaminyltransferase activity, called EryCIII.

8) Isolated DNA sequence represented in Figure 3 (sequence of SEQ ID No.6) corresponding to the *eryAI-eryK* region of the cluster of erythromycin biosynthesis genes.

9) DNA sequence according to claim 8 comprising:

- 10 - the *eryBIV* sequence corresponding to ORF13 (sequence of SEQ ID No.6 from nucleotide 242 to nucleotide 1207) and coding for a dTDP-4-keto-L-6-desoxyhexose 4-reductase,
- the *eryBV* sequence corresponding to ORF14 (sequence of SEQ ID No.6 from nucleotide 1210 to nucleotide 2454) and coding
- 15 for a mycarosyltransferase,
- the *eryCVI* sequence corresponding to ORF15 (sequence of SEQ ID No.6 from nucleotide 2510 to nucleotide 3220) and coding for a dTDP-D-6-desoxyhexose 3-N-methyltransferase,
- the *eryBVI* sequence corresponding to ORF16 (sequence of SEQ
- 20 ID No.6 from nucleotide 3308 to nucleotide 4837) and coding for a dTDP-4-keto-L-6-desoxyhexose 2,3-deshydratase,
- the *eryCIV* sequence corresponding to ORF17 (sequence of SEQ ID No.6 from nucleotide 4837 to nucleotide 6039) and coding for a dTDP-D-6-desoxyhexose 3,4-deshydratase,
- 25 - the *eryCV* sequence corresponding to ORF18 (sequence of SEQ ID No.6 from nucleotide 6080 to nucleotide 7546) and coding for a dTDP-D-4,6-didesoxyhexose 3,4-reductase and
- the *eryBVII* sequence corresponding to ORF19 (sequence of SEQ ID No.6 from nucleotide 7578 to nucleotide 8156) and
- 30 coding for a dTDP-4-keto-D-6-desoxyhexose 3,5 epimerase.

10) Isolated DNA sequence represented in Figure 3 chosen from the *eryBIV* sequence corresponding to ORF13 (sequence of SEQ ID No.6 from nucleotide 242 to nucleotide 1207), the *eryBV* sequence corresponding to ORF14 (sequence of SEQ ID No.6 from nucleotide 1210 to nucleotide 2454), the *eryCVI* sequence corresponding to ORF15 (sequence of SEQ ID No.6 from nucleotide 2510 to nucleotide 3220), the *eryBVI* sequence corresponding to ORF16 (sequence of SEQ ID No.6 from

09463705-032300

10

15

20

25

30

35

nucleotide 1207), eryBV (sequence of SEQ ID No.6 from nucleotide 1210 to nucleotide 2454), eryCVI (sequence of SEQ ID No.6 from nucleotide 2510 to nucleotide 3220), eryBVI (sequence of SEQ ID No.6 from nucleotide 3308 to nucleotide

4837), *eryCIV* (sequence of SEQ ID No.6 from nucleotide 4837 to nucleotide 6039), *eryCV* (sequence of SEQ ID No.6 from nucleotide 6080 to nucleotide 7546) or *eryBVII* (sequence of SEQ ID No.6 from nucleotide 7578 to nucleotide 8156)

5 represented in Figure 3, to synthesize hybrid secondary metabolites in *Sac. erythraea*.

16) Use of at least one of the DNA sequences chosen from the sequences *eryBII* (complementary sequence of SEQ ID No.1 from nucleotide 48 to nucleotide 1046), *eryCIII* (complementary
10 sequence of SEQ ID No.1 from nucleotide 1046 to nucleotide 2308) or *eryCII* (complementary sequence of SEQ ID No.1 from nucleotide 2322 to nucleotide 3404) represented in Figure 2, *eryBIV* (sequence of SEQ ID No.6 from nucleotide 242 to nucleotide 1207), *eryBV* (sequence of SEQ ID No.6 from
15 nucleotide 1210 to nucleotide 2454), *eryCVI* (sequence of SEQ ID No.6 from nucleotide 2510 to nucleotide 3220), *ebryBVI* (sequence of SEQ ID No.6 from nucleotide 3308 to nucleotide 4837), *eryCIV* (sequence of SEQ ID No.6 from nucleotide 4837 to nucleotide 6039), *eryCV* (sequence of SEQ ID No.6 from
20 nucleotide 6080 to nucleotide 7546) or *eryBVII* (sequence of SEQ ID No.6 from nucleotide 7578 to nucleotide 8156) represented in Figure 3 or of a fragment of this sequence, as hybridization probes.

17) Use of the *eryCIII* DNA sequence represented in Figure 2
25 (complementary sequence of SEQ ID No.1 from nucleotide 1046 to nucleotide 2308 = complementary sequence of SEQ ID No.4) as a hybridization probe to isolate genes responsible for the glycosylation of the macrolactone in a macrolide-producing strain.

30 **18)** Use according to claim 17 in which the homologous genes are the oleandomycin biosynthesis genes in *S. antibioticus*.

19) Isolated DNA sequence represented in Figure 22 (sequence of SEQ ID No.15) corresponding to a region of the cluster of oleandomycin biosynthesis genes comprising:

- 35 - the sequence corresponding to ORF *oleP1* from nucleotide 184 to nucleotide 1386,
- the sequence corresponding to ORF *oleG1* from nucleotide 1437 to nucleotide 2714 coding for a glycosyltransferase

- the sequence corresponding to ORF *oleG2* from nucleotide 2722 to nucleotide 3999 coding for a glycosyltransferase activity,

- 20) Isolated DNA sequence represented in Figure 22 chosen from the sequence corresponding to ORF *oleG1* (sequence of SEQ ID No.15 from nucleotide 1437 to nucleotide 2714 coding for a glycosyltransferase activity and the sequence corresponding to ORF *oleG2* (sequence of SEQ ID No.15 from nucleotide 2722 to nucleotide 3999) coding for a glycosyltransferase activity.

20 **22)** Isolated DNA sequence according to claim 20 corresponding to ORF oleG2 (sequence of SEQ ID No.15 from nucleotide 2722 to nucleotide 3999) coding for an oleandrosyltransferase activity.

24) Polypeptide coded by the DNA sequence corresponding to ORF *oleG2* and having an oleandrosyltransferase activity (sequence of SEQ ID No.18).

- a DNA sequence is isolated containing at least one *eryB* sequence or one *eryC* sequence of the cluster of erythromycin biosynthesis genes represented in Figure 2 (complementary sequence of SEQ ID No.1) or in Figure 3 (sequence of SEQ ID No.6),

- a modification is created in said sequence and an altered sequence is obtained,

- the altered sequence is integrated into the chromosome of the host strain and a modified strain is obtained,
- the modified strain is cultured under conditions which allow the formation of the hybrid secondary metabolite and
- 5 - the hybrid secondary metabolite is isolated.

26) Process according to claim 25 in which the DNA sequence codes for one of the enzymes chosen from a

- dTDP-4-keto-L-6-desoxyhexose 2,3-reductase,
- desosaminyltransferase,
- 10 - dTDP-4-keto-D-6-desoxyhexose 3,4-isomerase,
- dTDP-4-keto-L-6-desoxyhexose 4-reductase,
- mycarosyltransferase,
- dTDP-D-6-desoxyhexose 3-N-methyltransferase,
- dTDP-4-keto-L-6-desoxyhexose 2,3-deshydratase,
- 15 - dTDP-D-6-desoxyhexose 3,4-deshydratase,
- dTDP-D-4,6-didesoxyhexose 3,4-reductase or
- dTDP-4-keto-D-6-desoxyhexose 3,5 epimerase.

27) Process according to claim 25 in which the alteration of the sequence results in the inactivation of at least one of

20 the enzymes chosen from a

- dTDP-4-keto-L-6-desoxyhexose 2,3-reductase,
- desosaminyltransferase,
- dTDP-4-keto-D-6-desoxyhexose 3,4-isomerase,
- dTDP-4-keto-L-6-desoxyhexose 4-reductase,
- 25 - mycarosyltransferase,
- dTDP-D-6-desoxyhexose 3-N-methyltransferase,
- dTDP-4-keto-L-6-desoxyhexose 2,3-deshydratase,
- dTDP-D-6-desoxyhexose 3,4-deshydratase,
- dTDP-D-4,6-didesoxyhexose 3,4-reductase or
- 30 - dTDP-4-keto-D-6-desoxyhexose 3,5 epimerase.

28) Process according to claim 27 in which the inactivated enzyme is a dTDP-4-keto-L-6-desoxyhexose 4-reductase.

29) Process according to claim 27 in which the inactivated enzyme is a dTDP-D-6-desoxyhexose 3,4-deshydratase.

35 **30)** Process according to claim 27 in which the inactivated enzyme is a mycarosyltransferase.

31) Process according to claim 27 in which the inactivated enzyme is a dTDP-4-keto-L-6-desoxyhexose 2,3-reductase.

00220 5029450

32) Process according to claim 25 in which the isolated hybrid secondary metabolite is an analogue of erythromycin chosen from 4"-keto-erythromycin, 4'-hydroxy-erythromycin or 3"-C-desmethyl-2",3"-ene-erythromycin.

5 **33)** Process according to claim 25 in which the isolated hybrid secondary metabolite is desosaminy l erythronolide B.

34) Modified strain of *Sac. erythraea* in which at least one of the enzymes chosen from a

- dTDP-4-keto-L-6-desoxyhexose 2,3-reductase,
 - 10 - desosaminy l transferase,
 - dTDP-4-keto-D-6-desoxyhexose 3,4-isomerase,
 - dTDP-4-keto-L-6-desoxyhexose 4-reductase,
 - mycarosyltransferase,
 - dTDP-D-6-desoxyhexose 3-N-methyltransferase,
 - 15 - dTDP-4-keto-L-6-desoxyhexose 2,3-deshydratase,
 - dTDP-D-6-desoxyhexose 3,4-deshydratase,
 - dTDP-D-4,6-didesoxyhexose 3,4-reductase or
 - dTDP-4-keto-D-6-desoxyhexose 3,5 epimerase
- is inactivated and producing at least one hybrid secondary metabolite.

20 **35)** Modified strain of *Sac. erythraea* (BII92) in which a dTDP-4-keto-L-6-desoxyhexose 2,3-reductase is inactivated and producing 3"-C-desmethyl-2",3"-ene-erythromycin C.

36) Modified strain of *Sac. erythraea* (BIV87) in which a
25 dTDP-4-keto-L-6-desoxyhexose 4-reductase is inactivated and producing 4"-keto-erythromycin.

37) Modified strain of *Sac. erythraea* (CIV89) in which a dTDP-D-6-desoxyhexose 3,4-deshydratase is inactivated and producing 4'-hydroxyerythromycin D.

30 **38)** Modified strain of *Sac. erythraea* (BV88) in which a mycarosyltransferase is inactivated and producing desoaminy l erythronolide B.

39) Preparation process for precursors of oleandomycin in *S. antibioticus* in which

- 35 - an alteration is created in the sequence of the gene chosen from the DNA sequence corresponding to ORF *oleG1* (sequence of SEQ ID No.15 from nucleotide 1437 to nucleotide 2714) and the DNA sequence corresponding to ORF *oleG2* (sequence of

002220" 50/22460

SEQ ID No.15 from nucleotide 2722 to nucleotide 3999) in the chromosome of a host strain and a modified strain is obtained,

- the modified strain is cultured under conditions allowing the accumulation of the precursors of oleandomycin and
- these precursors are isolated.

40) Process according to claim 39 in which the alteration is created in the DNA sequence corresponding to ORF *oleG1* (sequence of SEQ ID No.15 from nucleotide 1437 to nucleotide 2714) and the result of which is at least the elimination of the desoaminytransferase activity and the accumulation of the precursor of oleandomycin 8,8a-desoxyoleandolide.

41) Thymidine 5'-(trihydrogen diphosphate), P'-[3,4,6-tridesoxy-3-(dimethylamino)-D-.xylo.-hexopyranosyl] ester (dTDP-D-desosamine) and the addition salts with bases.

006220 50269450

1/60

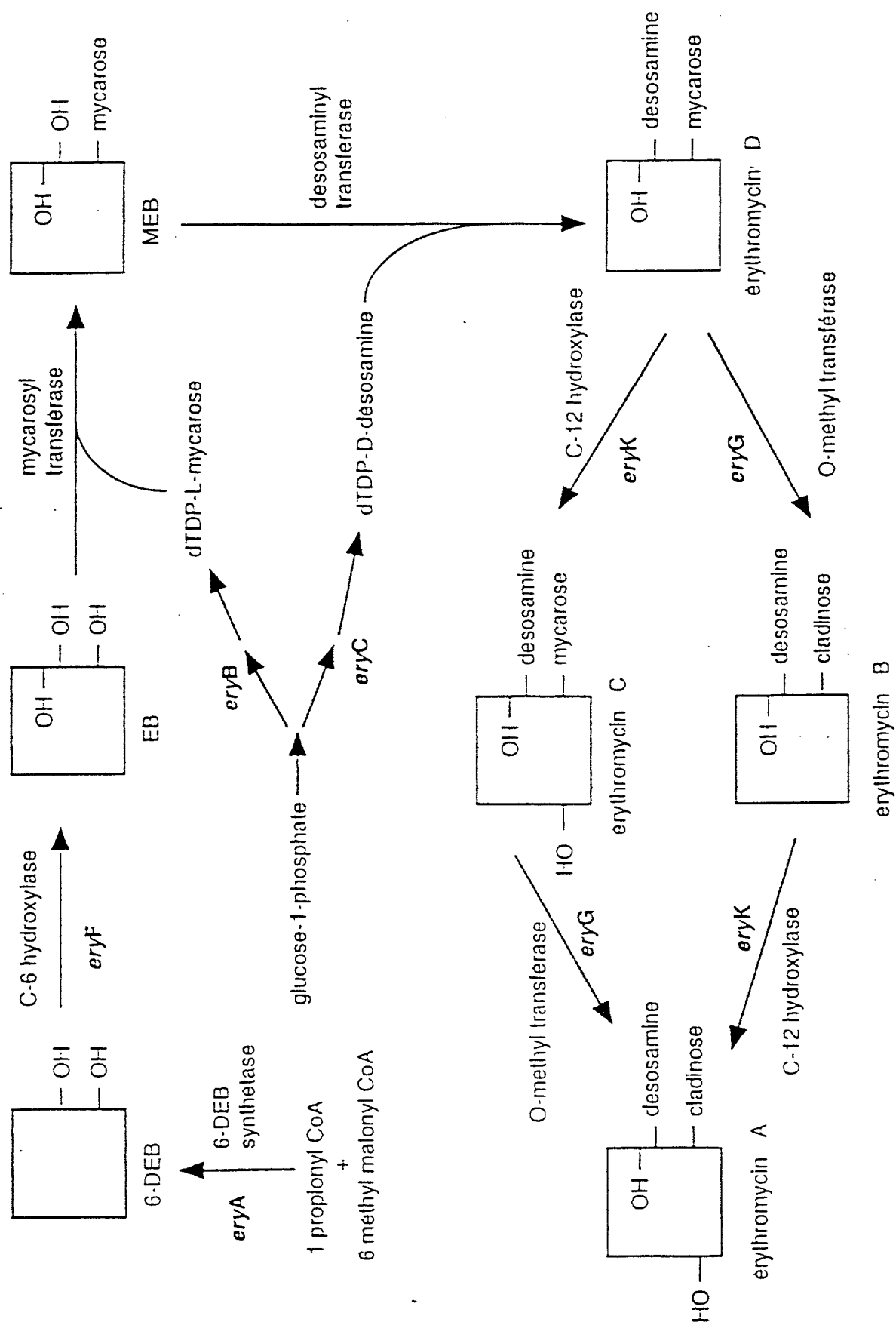


FIGURE 1

CTGCTTCACGCTCACCAGCCGTATCCITTC'CCGGTTCCTGTGCTCCTGCAACCAGG 7969
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
 GACGAAGTCCGAGTGGTCCGCATAGGAAAGAGCCCAAGGAGAACACAGGAGTGACGTTGGTCC
 Q K V S V * Q L W
 <---ORF6

 C'ITCCGGCGCCGCGCCGCGAGGCCACCGCGGGGAAGNTCTCGTCCAGTTCGGACAGCG 8029
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
 GAAGCGCGCGCGCGCGCGCTCCGGTGGCGCCCTTC'TAGAGCAGGTCAAGCCTGTTCGC
 A E P A A G G S A V A P F I E D L E S L

 CCTGCTCGTCCAGGGTCATCGCGGACGCC'ITCAGCGCGGAGTCGAGCTGCTCGGGGGTTC 8089
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
 GGACGAGCAGTCCAGTAGCGCTTCGGAAAGTCGGCCCTCAGCTCGACGAGCCCCCAAG
 A Q E D L T M A S A K L A S D L Q E P T

 GCGGCGCGATGACGGCGCGCGGATGCCGGCGCGGACAGCACCCATGCCAGCCCCACCT 8149
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
 CGCCCGCTACTGCCCGCGCGCGCTACGGCCCGCGCCCTGTGCTGGGTACGCTCGGGGTGGA
 R P G I V A G A I G P R S L V W A L G V

 CGCCCGGGTCT'ITCGCCGAGG'ITGCGGCAGAAC'ITCTGTAGGCCCTCGA'ITCGCCGGCGGCA 8209
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
 GCCGGCCCAAGAGCGCTCCCAACGCCGTCTTGAAGAGCATCCGGAGCTAGCGGCCCGCGGT
 E A P D E G L N R C F K E Y A E I A P R

 GGGACGGCAACAGCACCTTGGCACGGCCCTTGGCCCGAC'ITCACCGCGGTGCGCCCGCGGCA 8269
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
 CCC'TGCCGTGTGTGACGCGTCCCGGACGCGGCTGAAG'ITGGCGCCACGGGCGCGCGT
 L S P L L V Q A R G Q A S K V A T G A A

 GCTTCTCCAGCGCTCCGTGAGCAGGCCCGCGTGCAGCGCGGACCGAGCGGAGACGCCGA 8329
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
 CGAAGAGGTCCGAGGCGACTCGTCCGGGCGGACGTCGCCCGCTGGTCCGCTTCTGCGGCT
 L K E L A G S L L G G H L P S W A F V G

FIGURE 2

8330	GCCCGTAGGCC	TGCGCGCGGCGACAGCACAC	TCCAGC	TGGGC	G	TGGCGGACCGCCAGG	TGT	8389
	+	-----+	-----+	-----+	-----+	-----+	-----+	
	CGGGCATCCGGACGCGCGCCCG	TGCTGGAGAG	TGAGCGCGCACGGCC	TGGCGG	TCCAAACA			
	L G Y A Q A A A P L V E L E A H R V A L N							
8390	ACAGGCACTGGTGGAGACC	ATGCCACAGGGAG	TGGCGCGCGGCGCGG	TCTCTCTGCGCGG				8449
	+	-----+	-----+	-----+	-----+	-----+	-----+	
	TGTCCGTGACCACCC	TCTGGTACGGGT	TCCC	TACCGCGCGCGCGCAAGAGGACGCGCC				
	Y L C Q H S V M G L S H R R A A N E Q A							
8450	CGCGGATGTGCCAGCCCGCGAAG	TTCGACGAGCGCAGCTAGGAGACC	TGCGCGCTGGCGGA					8509
	+	-----+	-----+	-----+	-----+	-----+	-----+	
	GCCGCTACACGGT	CGGCGCTCAAGC	TGCTGGCTGCATCCT	TGGAACGGCGACCGCT				
	A A I H W G A F N S S G V Y S V K G S A							
8510	CGAGGCTGTCCCATGGCC	TGCCACACC	TGCTCCACCGCGGACCGG	TGATGTGGTGC				8569
	+	-----+	-----+	-----+	-----+	-----+	-----+	
	GCTCCGACAGGTACCGGACGGT	TGTGGAGCAGGGT	TCCCGCGCC	TGGCCAGCTACACCACGT				
	V L S D M A Q W V E D W P A S R D I H H							
8570	TCTGGTAGACG	TGATGTGGTTCGACGCGCCAGCC	TGCGCAGCGATCCCT	TGCGAGGAGCGGA				8629
	+	-----+	-----+	-----+	-----+	-----+	-----+	
	AGACCATCTGCAGCTACACCAGC	TGCGGGTTCGACCGCTCGCTAGGGAGCGTCT	TCCGCT					
	M Q Y V D I H D V G L R R L S G E C S A							
8630	TGATGTGCCCCGCGCCGACAGCCCCG	TGTCTGT	TGACCGCGT	TGCGTCA	TCTGCGCGCGCGACCT			8689
	+	-----+	-----+	-----+	-----+	-----+	-----+	
	ACTACACGCGCGGCTGT	CGGGCGACAGCAACT	TGCGGAGCGAGTAGAGCGCGGCT	TGGA				
	I I H R A S L G S D N V R E S M E G G V							
8690	TGGTCGCCAGCACGGTGT	CTCGCGCGCT	TCCGCGCGCCCT	TGGCGCAGCCACCT	TGCCCCACCA			8749
	+	-----+	-----+	-----+	-----+	-----+	-----+	
	ACCAGCGGTCTGTGCCACAGGACGCGGCGGACCGG	TGCGGTGGACGGGT	TGGT					
	K T A L V T D E R R G G G Q A L W R G V							

```

8750  GC'TCC'TCGGTG'TGGCCCTTGT'GAGAGCCGCCAGCC'G'PACATG'TCGGCGGTG'TCGAGGCAGT
      +-----+-----+-----+-----+-----+-----+-----+-----+
      CGAGGAGCCACACGGGAACAT'CTCGGCGGTGCGCATGT'ACAGCCGCCACAGCT'CCGTCA
      L E E T H G K Y L R W G Y M D A T D L C
8809
8810  TGATGCCGGGTCCCGGGGTGGTCCATCAGGGCGCAGCGGTG'TCG'TCC'TCGACGCGTC
      +-----+-----+-----+-----+-----+-----+-----+-----+
      ACTACGGCGCCAGGGCCCGCACCAGG'PAG'TCCGCGTTCGGCAGCAGCAGGAGCTGCGCAG
      N I G R D R A H D M L R L A D D D E V R
8869
8870  CGCTGAAGTTCACCGTCCCGAGCCAGAGCCTGCTGG'TGAGCAGCGCGGAACGCCGAGCC
      +-----+-----+-----+-----+-----+-----+-----+-----+
      GCGACTTCAAGTGGCACGGCTCGGTC'TCGGACGACCAC'TCGTTCGGCGCTTTCGGGCTCGG
      G S F N V T G L W L R S T L L A S R G L
8929
8930  GCACG'TCGGTGCGGGGTGGTGG'ICATCGTGGT'IC'ICCT'ICCT'GCGGCCAGTTCCTC
      +-----+-----+-----+-----+-----+-----+-----+-----+
      CGTGACGCGCGCGCAGCCACCAGTAGCACCAAGAGAGGAAGGACGCCGGTCAAGGAG
      R V H T A A D T T M
      <---ORF7 * R P E R R G A A L E E
8989
8990  GCAGATGCCGACGACCTCGGCCGGTGACGGCTCCCGGAGCATG'TCG'TCGCGCATCCCGCC
      +-----+-----+-----+-----+-----+-----+-----+-----+
      CGTCTACGGCTGCTGGAGCCGGCCAC'TGCCGAGCGCGTCTG'TACAGCAGCGCGTAGGGCGG
      C I G V V E A P S P E A L M D D R M R A
9049
9050  CGCGCCGGCGGGTGGGCCGGGTG'TCGAGGACCCGCT'ICACCGACTCCCGGAGCTGGTC
      +-----+-----+-----+-----+-----+-----+-----+-----+
      GCGGGCGCGGCCACCCGGCCAGCAGCTCC'TGGCGGAAGTGGCTGAGGGGCC'ICGACCCAG
      A G A R H A P D D L V R K V S E R L Q D
9109
9110  GGGGGTCAGCTCGGGCACGGGCAGCGCGATCCCCGCCCGGAA'TTCTGCGTGGCGTGGCGC
      +-----+-----+-----+-----+-----+-----+-----+-----+
      CCCCCAGTCGAGCCCGTGCCCGTTCGCGCTAGGGCGGGGCT'PAAGGACGACGCGACGCG
      P T L E P V P L A I G A G F E Q T R Q A
9169

```

FIGURE 2

9170 GCGACGCCCGGTGTCCCAAGCCGTCCGGCAGGATCACCTGCGGCACCCCGTGGATTCGCCCGC
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
 CGCGTGGGGCCACAGGGTCCGGCAGCCCGTCCCTAGTGGACGCCGTGCGGCACCTAGCGGCG
 R V G T D W G D P L I V Q P V G H I A A 9229
 GGTGTGCCAGCTCCCGGGTCCCGCCGTGGTGCACCGTCCGCCGCGCAGGTCCGGCAGCAGCGC 9289
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
 CCACACGGTCCGAGGGCCCGCCAGCGGCACACAGTGGCAGCGCGCGTCCAGCCGTCTGTCGCG
 T H W S G P G G H H V T A A C T P L L A
 GTGCATCGGGACGAAGCCGACCGTCCGGACGTTCGCGGGATGTTCGCGGACGCCCTTCTAG 9349
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
 CACGTAGCCCCTGCCTTCGGCTGGCACGCCCTGCACAGGCCCTACAGCGCTGCGGAAGATC
 H M P V F G V T R V N D P I N A V G E L
 CTTCTGCGCGTCGAAGGTCCGGATGATCTCGGCCTCGACGTTCGCCGACGGCACCCAGCAG 9409
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
 GACGACGCGCAGCTTCCAGCGCTACTAGAGCCGCGAGCTGCAGCGGTGCCGTGGGTCTGTC
 Q Q A D F T A I I E A D V D G V A G L L
 CTCTCGATGGAGACCTGCCCGGATGCTGTCTTCGCGCGCTGGAGATCCCGAGCGTGGAGGCA 9469
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
 GAGGAGCTACCTCTGGACGGGTACGACAGAGCGCCGACCTCTAGGGCTCGCACCTCCGT
 E E I S V Q G I S N E R S S I G L T L C
 CACGCGCGCGCGCTCCGGCTCGTTCGTGACGCCATTCGGGCACCCAGGACGGCCCGTGTGA 9529
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
 GTGCGCCGCGCGAGCCCGAGCAGCACGTCCGTAAAGCCGTGGTGGCTGCGCGGCAACAT
 V R R R E P E D H L W E P V V S P G N Y
 GTTCGACGTAGCGCATCCCGACGGTCTTCAGGCCGGTGTTCGAGCCCTGATCGCGGCCCGGGC 9589
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
 CAGCTGCATCGCGTAGGGCTGCCAGAAGTTCGGGCCACAGCTCGGACTAGCGCCGGCCCGG
 D V Y R M G V T K L G T D L R I A A P A

FIGURE 2

9590	GGGTCGATCGTCCACATGCCCCGACGACCACCTCCTCGTGAAAGCCCGGCCCGTACTT	9649
	CCCCAGCTAGCAGGTGACGGGCTGCTGGTGGAGGAGCAGCTTCCGGCCCGCGGCATGAA	
	P D I T W Q G V V E E D F A P G G Y K	
9650	CTCCAGCGTCCAGGTGAGCCACTCGGCGAGCGGTCTCCCGGTGCTCCTCCGGCTGGTCTC	9709
	GAGGTGCGAGGTCCACTCGGTGAGCCGCCTCGCCACGAGGCGCCACGAGGAGGCCGACCAG	
	E L T W T L W E A L P D E R H E E P Q D	
9710	GGCAGCAGGCGGAGGAAGTTCTGCCCCGCCCGGGTGGTGAATGTCGGGTCCCCACAGCAG	9769
	CCCCGTGTCGGCTCCCTTCAAGACGGCGCGGGCCCCACCACCTACAGCCCCAGGGTGTCTGTC	
	P L L G L F N Q R A R T T I D P G W L L	
9770	CCGCGGTGCGGGCTTCCGGTCAACGCGCGCGCGAATGGCGCGCGGAAGGTGAGCGGCTC	9829
	GGCGGCACGCGCAAGGCCAGTGGCGGCGGCGCTTACCGCGCGCGCTTCCACTCGCCGAG	
	R A H P T G T V A A A I P A A F T L P E	
9830	CCAGATGACCAGGTGCGGGCGGCCACTTCCGGCAGAACGAGACCATGCCCTTCGATGAGCGT	9889
	GGTCTACTGGTCCAGCCCCGGGTGAAGCGCTTGTGCTCTGGTACGGAAGCTACTCGCA	
	W I V L D P R W K R C F S V M G E I L T	
9890	GTCCGGGCTCATCAGGGCGTAGAAGGTGGGGTGGAGCACGGTCTGCATGCCCCAGCAGGTG	9949
	CAGGCCCCGAGTAGTCCCGCATCTTCCAGCCCCACCTCGTGGCCAGACGTACGGGTGCTCCAC	
	D P S M L A Y F T P T L V T Q M G L L H	
9950	CTCCCAGGTCAAGGTGGCGGGTCCCGCTCGCTGAAGTCCAGGCTCCGGACGTAGTCTGAT	10009
	GAGGGTCCAGTTCCACCGCCCCCAGGCGGAGCGACTTTCAGGTCCGAGGCCCTGCATCAGCTA	
	E W T L T A P D R E S F D L S R V Y D I	

FIGURE 2

10010 GATGTCGTGGCCCCGGTGGGTCAVGAAGTCCACGAGGTCGACGTCGGTGCCGACCCGGGAC 10069
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
 C'TACAGCACCGGGCGCACCCAGTACTTCAAGGTGCTCCAGCTGCAGCCACGGCTGGCCCTG
 I D H G A H T M F D V L D V D T G V P V
 10070 GCGGGTCAGCCCCGGCGGGTGATGTCTCGGTGAGCGCCGGGACGCGACCCACGCGGAC 10129
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
 CCGCCAGTCGGGCGGCGCCACTACAGGAGCCACTCGCGGCCCTTGGCGTGGTGCGCCTG
 A T L G A A T I D E T L A P S A V V R V
 10130 C'TCGTGCCCCCGCGCGGAACGCCCATGCGAGGGGACGAGGCCGGAAGAGGTGGCTCTT 10189
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
 GAGCACGGGGCGGCGGCCCTTGGCGGTACGCTCCCCCTGC'TCCGGGTCT'TCCACCGAGAA
 E H G A A R F A W A L P V L G F L H S K
 10190 GCTGGCCATGAGGAGAAGACGACGCGCATCGCGGTACCTACAGAGCTCGACGGGGCAGC 10249
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
 CGACCGGTACCTCTTCTGCTGCGGTAGCGCCCAATGGAGTCTCGAGCTGCCCCCGTCG
 S A M S S F V V R M * L E V P C
 <---ORF8
 10250 GGT'TGGT'TCCCCGAGGACGGGTGATCGGGCGCGCCGACGACCGGGCCGCTGGGGCGTGA 10309
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
 CCAACCAAGGGCGTCTTGGCCACTAGCCCGCGCGGCGCTGC'TGGCCCCGGCGACCCCGCACT
 R N T G R L V P S R R R R V V P G S P T
 10310 GTCCGGGACGCGCTTGGCCCGCGCCCGCAGTCCGGGGGTGGCGAGCGCGT'GACCAGCT' 10369
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
 CAGGCCCGTTCGGGAACCGCGCGCGGCGTACGCCCGCCACCGCTCGCGCCACTGGTCGA
 L G P L A K A A A R L A A T A L A T V L
 10370 CCTCCAGCC'TGCCGGGGTGGCCGGGATGTGCCGACAGCGCGCGGTCCGGCTCGGGCGCGGT 10429
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
 GGAGGTCCGACGGCCCCACCGGGCGTACACGGCTG'TCGCGCGCCAGCCGACGCCCGCCCA
 E E L R G P H G R H A S L A R D A D P R

FIGURE 2

10430 CCACGTCGAGGGGTCGGGCTCGGCGAAGACCTCCGGGTCGGGTTCGGCCGCGGACGA 10489
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
GGTGACGCTCGGCCAGCCGAGCCGCTTCGAGAGCCAGCCGCAACCCGGCGGCGTGCCT
D V D L R D P E A F V E P D R N A A V
10490 CGACCACGACCTCCTCGCCCTTCGCCGATCACGTGCTCGCCGAGCCGCACCTCTGCGGTGG 10549
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
GCTGTGCTGGAGGAGCGGAAGCGGCTAGTCACAGAGCGGCTCGGCGTGGAGACGCCACC
V V V V E E G E G I V H E G L R V E A T
10550 CCGTGGCGCGCTCCAGGTGCAATGCCCGGTGCAGGCGCAGCACCTCCGGGACGGTTCGCT 10609
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
GGCACGCGCGGAGGTCCACGTTCACGGCCACAGTCGCGGTTCGTCGAGCCGCTGCCCAAGCGA
A T R R E L H L A P H L R L V E A V T R
10610 GCGCGCGCGGGTCTGTCGGCGATCCGTCGGCCAGCCCCCGTTCGGCCGAGACGGCCA 10669
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
CGCGCGCGCGCCAGCAGCCGCTAGGCAAGCCGTCGGGGCCCAAGCCGCTTCGCGCGT
Q A A A P D A I R E A L G P E A S V A
10670 GGACCGGTCGACACCGGTCTCGCGGTCACTTCGGCCCCCGCGAACAGGGCGCGCAGTG 10729
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
CCTGGCGCAGTGGTCCACACAGCGCCAGTAGAGCCGGCGCGCTTGTCGCGCGCTCAC
L V A D V V T N A T M E A G A F L A R L
10730 CGGGTTCGGCGGCGAGTCCCCGCGACCGCTGCTTCGGTTCACCGCGAGCTGCTGCGGGCTGA 10789
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
GCCCGACCGCGCGTACCGCGCTGGCGACGAGCCAGTCGGCGCTCGACGACGCCCGGACT
A P D A P L A A V A A E T V A L Q Q P S
10790 GCTGGCGCTCCAGGCTGACGCGGGCGTCCACAGCGCGCGCGCGCAGCACTCCGGCTGCGC 10849
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
CGACCCCGCAGGTCCGACTTCGCGCGCGCAGGGTTCGCGCGCGCGCGCTGTCGTCGAGCCGCGCGG
L Q A D L S V R A D W A A G R L V G A A

FIGURE 2

CGAGCACGGCGGTCAATGCCCTGCACCGGTATACCTTGCCAGGCGAAGTTCGCCGACCGACGAGTCCA
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
GCTCGTCCGCCAGTACGGGACGTGGCCAATGGACGGTCCGCTTCAGCGGGCTGGTCCAGGT
G L V A T M G Q V P V Q W A F D G V L D
10850

GCGCGCGCCCGCGCGGGGAGCAGACCGCGGAGAGCTCTCCGCCAGTTCCCGCAGCGTCGG
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
CGCGCGCGGGCGGCCCTCGTCTGGCCGCTTCGAGAGGGCGGTCAAGGGGCTGCAGCC
L R A G A G P L L G A F S E A L E G V D
10910

GGACCTCGCCCTTCCAGGACGCGGCGTGCACGTCCCGGAACGGCTGGGCCACCTCGGGCGG
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
CCTGGAGCGGAAGGTCTTCGCGCGCACGATGCAGGGCTTTCGCCGACCGGGGTGAGCGCGCC
P V E G E W S A A H V D R F P Q A W E A
10970

GTGGCGCGCCCGCGCGCATCCAATTCGGGTGTCGGTCCGGTGGCGCGGGTGAACGCGG
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
CACCGCGGGCGCGGGCTAGGTAAAGGCCACACGAGGCCACCGCGGCCACTTGCGCC
P P A G A A R M W E P T R G T A R T F A
11030

GGTCTGTGAGCACCTTCGGCGCGGTGGCTGGTGGCCACCAACCCACGTCTCGGGTGGCGG
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
CCAGCAGCTCGTGGACGGCGCGCCACCGCACCGCGGTGGTGGTTCAGAGCGCCACGCGG
P D D L V Q R A T A H D A V V W T E T R
11090

TGCGCGGCACACCGGACTTCGGGCATCGAGCGGTACCGCGCTGGCGGTCTGTCGTCTGTCTC
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
ACCGGGCGTGTGGCTTGAGCGCGTAGCTCGCCCATGGCCCGGACGCCCGCAGCAGCACACAG
S R R V G S E R M S R Y R R Q P D D H
CGCACAGCAGCATCGGGTAAGGGTCGCCGTTCGTGCCGTAAACCCACAGTGCAGGCGCGCGGA
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
GCGTGTCTGTAGCCCATTCGCCAGCGGCACGACGGCATTTGGGGTCACTCCGGCGCCT
G C L L M P Y P D G N S G Y G W H L G R
11210

FIGURE 2

00E220" 50/E 2460

WO 99/05283

09/463705
PCT/FR98/01593

10/60

```
11270 TCAATCTGGAGCTGCCCTGCCCCAGCCCCGGCGCGATCGGTCTGGTGCATGAATTC 11321
+-----+-----+-----+-----+-----+-----+-----+-----+
AGTAGACCTCGACGGACGGGTCCGGCCCGCGCTAGCCAGCACCCAGTACTTAAG
I M Q L Q R G L G A R D T T M
<---ORF9
* S N
<---ORF10 ...
```

FIGURE 2

11/60

43610 43630 43650
TTTGACAGGTCCGCCACGCGTCCCCCTACTCGACGACCACGCAATGGGCGAACAATATAG
43670 43690 43710
GAAGGATCAAGAGGTTGACATCGCCTCGTCGAGCCACGAACCTGTGAACATCTGCATGT
43730 43750 43770
TGACAAGATCAACGGCGGCTACCTACTGTGGTGGCCCCAGTGACGGGTTGCCGCACATCGC
43790 43810 43830
GCTGGGGAGATTCTTTGAATTTGCCCCGTAGCACCGACCTGGAAAGCGAGCAAATGCTCC
43850 43870 43890
GGTGAATGGGATCAGTGATTCCCCGCGTCAATTGATCACCTTCTGGGCGCTTCCGGCTT
V N G I S D S P R Q L I T L L G A S G F
ORF13 --->
43910 43930 43950
CGTCGGGAGCGCGGTTCTGCGCGAGCTGCGCGACCACCCGGTCCGGCTGCGCGCGGTGTC
V G S A V L R E L R D H P V R L R A V S
43970 43990 44010
CCGCGGCGGAGCGCCCCGCGGTTCCGCCCCGGCGCCGCGGAGGTCGAGGACCTGCGCGCCGA
R G G A P A V P P G A A E V E D L R A D
44030 44050 44070
CCTGCTGGAACCGGGCCGGCCGCGCCGCGATCGAGGACGCGGACGTGATCGTGACCT
L L E P G R A A A A I E D A D V I V H L
44090 44110 44130
GGTGGCGCACGCAGCGGGCGGTTCCACCTGGCGCAGCGCCACCTCCGACCCGGAAGCCGA
V A H A A G G S T W R S A T S D P E A E
44150 44170 44190
GCGGCTCAACGTCGGCCTGATGCACGACCTCGTCGGCGCGCTGCACGATCGCCGCAGGTC
R V N V G L M H D L V G A L H D R R R S
44210 44230 44250
GACGCGCGCCCGTGTGCTCTACGCGAGCACCGCACAGGCGCGAACCCTCGCGCGGCCAG
T P P V L L Y A S T A Q A A N P S A A S
44270 44290 44310
CAGGTACGCGCAGCAGAAGACCGAGGCCGAGCGCATCCTGCGCAAAGCCACCGACGAGGG
R Y A Q Q K T E A E R I L R K A T D E G
44330 44350 44370
CCGGGTGCGCGGCGTGATCCTGCGGCTGCCCGCGCTCTACGGCCAGAGCGGCCCGTCCGG
R V R G V I L R L P A V Y G Q S G P S G
44390 44410 44430
CCCCATGGGGCGGGGCGTGGTCGCAGCGATGATCCGGCGTGGCCCTCGCCGCGAGCCGCT
P M G R G V V A A M I R R A L A G E P L
44450 44470 44490
CACCATGTGGCACGACGGCGGCGTGCGCCGCGACCTGCTGCACSTCGAGGACGTGGCCAC
T M W H D G G V R R D L L H V E D V A T
44510 44530 44550
CGCGTTCCGCCCGCGCGCTGGAGCACCACGACGCGCTGGCGCGGCGCACGTGGGCGCTGGG
A F A A A L E H H D A L A G G T W A L G
44570 44590 44610
CGCCGACCGATCCGAGCCGCTCGGCGACATCTTCCGGGCGGCTCTCCGGCAGCGTCGCCCCG
A D R S E P L G D I F R A V S G S V A R
44630 44650 44670
GCAGACCGGCAGCCCCGCGCTCGACGTGGTCACCGTGCCCCGCGCCCGAGCACGCGCGAGGC
Q T G S P A V D V V T V P A P E H A E A

FIGURE 3

12/60

44690 44710 44730
 CAACGACTTCCGCGACGACATCGACTCCACCGAGTTCCGCGAGCCGGACCGGCTGGCG
 N D F R S D D I D S T E F R S R T G W R
 44750 44770 44790
 CCCCCGGGTTTCCCTCACCAGCGGCATCGACCGGACGGTGGCCGCCCTGACCCCCACCGA
 P R V S L T D G I D R T V A A L T P T E
 44810 44830 44850
 GGAGCACTAGTGCGGGTACTGCTGACGTCCTTCGCGCACCGCACGCACTTCCAGGGACTG
 E H *
 V R V L L T S F A H R T H F Q G L
 ORF14 --->
 44870 44890 44910
 GTCCCGCTGGCGTGGGCGCTGCGCACCGCGGGTCACGACGTGCGCGTGGCCGCCAGCCC
 V P L A W A L R T A G H D V R V A A Q P
 44930 44950 44970
 GCGCTCACCGACGCGGTTCATCGGCGCGGTCCTCACCAGCGGTACCCGTCGGCTCCGACCAC
 A L T D A V I G A G L T A V P V G S D H
 44990 45010 45030
 CGGCTGTTTCGACATCGTCCCGGAAGTCGCGGCTCAGGTGCACCGCTACTCCTTCTACCTG
 R L F D I V P E V A A Q V H R Y S F Y L
 45050 45070 45090
 GACTTCTACCGCGGAGCAGGAGCTGCACTCGTGGGAGTTCTGCTCGGCATGCAGGAG
 D F Y H R E Q E L H S W E F L L G M Q E
 45110 45130 45150
 GCCACCTCGCGGTGGGTATACCGGTGGTCAACAACGACTCCTTCGTCGCCGAGCTGGTC
 A T S R W V Y P V V N N D S F V A E L V
 45170 45190 45210
 GACTTCGCCCGGGACTGGCGTCTGACCTGGTGTCTGGGAGCCGTTACCTTCGCCCGGC
 D F A R D W R P D L V L W E P F T F A G
 45230 45250 45270
 GCCGTGCGGCCCCGGGCTGCGGAGCCGCGCACGCCCGGCTGCTGTGGGGCAGCGACCTC
 A V A A R A C G A A H A R L L W G S D L
 45290 45310 45330
 ACCGGCTACTTCCGCGCGCGGTTCCAGGCGCAACGCCTGCGACGGCCGCGGAGGACCGG
 T G Y F R G R F Q A Q R L R R P P E D R
 45350 45370 45390
 CCGGACCCGCTGGGCGACGTGGCTGACCGAGGTGCGGGGGCGCTTCGGCGTGAATTCCGGC
 P D P L G T W L T E V A G R F G V E F G
 45410 45430 45450
 GAGGACCTCGCGGTGCGGCGAGTGGTGGTTCGACCGAGTTGCCGCGGAGTTTCCGGCTGGAC
 E D L A V G Q W S V D Q L P P S F R L D
 45470 45490 45510
 ACCGGAATGGAAACCGTTGTGCGCGGACCCCTGCCCTACAACGGCGCGTGGTGGTTCCG
 T G M E T V V A R T L P Y N G A S V V P
 45530 45550 45570
 GACTGGCTCAAGAAGGGCAGTGGGACTCGACGCATCTGCATTACCGGAGGGTTCTCCGGA
 D W L K K G S A T R R I C I T G G F S G
 45590 45610 45630
 CTCGGGCTCGCCGCCGATGCCGATCAGTTCGCGCGGACGCTCGCGCAGCTCGCGCGATTG
 L G L A A D A D Q F A R T L A Q L A R F
 45650 45670 45690
 GATGGCGAAATCGTGGTTACGGGTTCCGGTCCGGATACCTCCGCGGTACCGGACAACATT
 D G E I V V T G S G P D T S A V P D N I

FIGURE 3

00463705 50463705

13/60

45710 45730 45750
CGTTTGGTGGATTTCGTTCCGATGGGCGTTCTGCTCCAGAACTGCGCGGCGATCATCCAC
R L V D F V P M G V L L Q N C A A I I H
45770 45790 45810
CACGGCGGGGCGGAACCTGGGCCACGGCACTGCACCACGGAATTCCGCAAATATCAGTT
H G G A G T W A T A L H H G I P Q I S V
45830 45850 45870
GCACATGAATGGGATTGCATGCTACGCGGCCAGCAGACCGCGGAAGTGGGCGCGGGAATC
A H E W D C M L R G Q Q T A E L G A G I
45890 45910 45930
TACCTCCGGCGGACGAGGTGCGATGCGGACTCATTTGGCGAGCGCCCTCAGCCAGGTGGTC
Y L R P D E V D A D S L A S A L T Q V V
45950 45970 45990
GAGGACCCACCTACACCGAGAACGCGGTGAAGCTTCGCGAGGAGGCGCTGTCCGACCCG
E D P T Y T E N A V K L R E E A L S D P
46010 46030 46050
ACGCCGCGAGGAGATCGTCCCGCGACTGGAGGAACCTCAGCGCGCCGCCACGCGCGCTAGCGG
T P Q E I V P R L E E L T R R H A G *
46070 46090 46110
TTTCCGACCGACAAGTCCGTCCGACAGCACACCTCCGGAGGGAGCAGGGATGTACGAGGG
M Y E G
ORF15 --->
46130 46150 46170
CGGGTTCGCCGAGCTTTACGACCGGTTCTACCGCGGCGGGGCAAGGACTACGCGGCCGA
G F A E L Y D R F Y R G R G K D Y A A E
46190 46210 46230
GGCCGCGCAGGTGCGCGGGCTGGTCAGAGACCGCCTGCCCTCGGCTTCCTCGCTGCTCGA
A A Q V A R L V R D R L P S A S S L L D
46250 46270 46290
CGTGGCCTGCGGGACCGGCACCCACCTGCGCGGTTTCGCCGACCTCTTCGACGACGTGAC
V A C G T G T H L R R F A D L F D D V T
46310 46330 46350
CGGGCTGGAGCTGTGCGCGCGGATGATCGAGGTGCGCCCGGCCGACGCTCGGCGGCATCCC
G L E L S A A M I E V A R P Q L G G I P
46370 46390 46410
GGTGCTGCAGGGCGACATGCGCGACTTCGCGCTGGATCGCGAGTTTCGACGCCGTACCTG
V L Q G D M R D F A L D R E F D A V T C
46430 46450 46470
CATGTTTCAGCTCCATCGGGCACATGCGCGACGGCGCCGAGCTGGACCAGGCGCTGGCGTC
M F S S I G H M R D G A E L D Q A L A S
46490 46510 46530
CTTCGCCCCGCCACCTCGCCCCCGGCGCGTCTGTGGTGGTGAACCGTGGTGGTTCCCGGA
F A R H L A P G G V V V V E P W F P E
46550 46570 46590
GGACTTCCTCGACGGCTACGTGGCCGCGTACGTGGTGGCGGACGGCGACCTGACGATCTC
D F L D G Y V A G D V V R D G D L T I S
46610 46630 46650
GCGCGTCTCGCACTCCGTGCGCGCGCGGCGGCGGACCCGGATGGAGATCCACTGGGTGCT
R V S H S V R A G G A T R M E I H W V V
46670 46690 46710
GGCCGACGCGGTGAACGGTCCGCGGCAACACGTGGAGCACTACGAGATCAGCTCTTCGA
A D A V N G P R H H V E H Y E I T L F E

FIGURE 3

14/60

46730 46750 46770
GCGGCAGCAGTACGAGAAGGCCTTCACCGCGCGCGGTTGCGCTGTGCAGTACCTGGAGGG
R Q Q Y E K A F T A A G C A V Q Y L E G
46790 46810 46830
CGGACCCTCCGGACGCGGGTGTTCGTCGGTGTGCGCGGATGACCCGTGCGTTCGCGTTT
G P S G R G L F V G V R G *
46850 46870 46890
TCCGTTCCTGGCACAGGTGATCCGCTCCACGGGCCCTTTCCCCGCCGTGACCGGACCCTT
46910 46930 46950
ACAGTGAGTGCGGGTCTTGATCGACAACGCCCCGGCGGCAGCAAGCGGAGCCGTGACGAC
V R V L I D N A R R Q Q A E P S T T
ORF16 --->
46970 46990 47010
ACCGCAGGGAGAGTGCATGGGTGATCGGACCGGCGACCGGACGATTCCCGAATCCTCGCA
P Q G E S M G D R T G D R T I P E S S Q
47030 47050 47070
GACCGCAACGCGTTTCTCTGCTCGGCGACGGCGGAATCCCCACCGCCACGCGGAAACCCA
T A T R F L L G D G G I P T A T A E T H
47090 47110 47130
CGACTGGCTGACCCGCAACGGCGCCGAGCAGCGGCTCGAGGTGGCGCGCGTGGCGTTGAG
D W L T R N G A E Q R L E V A R V P F S
47150 47170 47190
CGCCATGGACCGCTGGTCTCCAGCCCCGAGGACGGCAGGCTCGCCCCACGAGTCCGGGCG
A M D R W S F Q P E D G R L A H E S G R
47210 47230 47250
CTTCTTCTCCATCGAGGGCCTGCACGTGCGGACGAACCTTCGGCTGGCGGCGGGACTGGAT
F F S I E G L H V R T N F G W R R D W I
47270 47290 47310
CCAGCCCATCATCGTGCAGCCCCGAGATCGGCTTCCTCGGCCTCATCGTCAAGGAGTTCA
Q P I I V Q P E I G F L G L I V K E F D
47330 47350 47370
CGGTGTGCTGCACGTGCTGGCGCAGGCCAAGGCCGAGCCGGGCAACATCAACGCCGTCCA
G V L H V L A Q A K A E P G N I N A V Q
47390 47410 47430
GCTCTCCCCGACCCTGCAGGCGACCCGACGCAACTACACCGGCGTCCACCGCGGCTCGAA
L S P T L Q A T R S N Y T G V H R G S K
47450 47470 47490
GGTCCGGTTCATCGAGTACTTCAACGGCACGCGCCCCGAGCCGGATCCTCGTGCAGCTGCT
V R F I E Y F N G T R P S R I L V D V L
47510 47530 47550
CCAGTCCGAGCAGGGCGCGTGGTTCCTGCGCAAGCGCAACCGGAACATGGTTCGAGGT
Q S E Q G A W F L R K R N R N M V V E V
47570 47590 47610
GTTTCGACGACCTGCCCCGAGCACCCGAACCTTCGGTGGCTGACCGTCGCGCAGCTGCGGGC
F D D L P E H P N F R W L T V A Q L R A
47630 47650 47670
GATGCTGCACCACGACAACGTGGTGAACATGGACCTGCGCACCGTGCTGGCCTGCGTCCC
M L H H D N V V N M D L R T V L A C V P
47690 47710 47730
GACCGCGGTGGAGCGGGACCGGGCGGACGACGTGCTCGCGCGCCTGCCCGAGGGCTCGTT
T A V E R D R A D D V L A R L P E G S F

FIGURE 3

15/60

47750 47770 47790
 CCAGGCCCGGCTGCTGCACTCGTTTCATCGGCGCGGGCACCCCGGCCAACATGAACAG
 Q A R L L H S F I G A G T P A N N M N S
 47810 47830 47850
 CCTGCTGAGCTGGATCTCCGACGTGCGCGCCAGGCGCGAGTTCGTGCAGCGCGGCCGCC
 L L S W I S D V R A R R E F V Q R G R P
 47870 47890 47910
 GCTGCCCCGACATCGAGCGCAGCGGGTGGATCCGCCGCGACGACGGCATCGAGCACGAGGA
 L P D I E R S G W I R R D D G I E H E E
 47930 47950 47970
 GAAGAAGTACTTCGACGTCTTCGGCGTCACGGTGGCGACCAGCGACCGCGAGGTCAACTC
 K K Y F D V F G V T V A T S D R E V N S
 47990 48010 48030
 GTGGATGCAGCCGCTGCTCTCGCCCGCCAAACAACGGCCTGCTCGCCCTGCTGGTCAAGGA
 W M Q P L L S P A N N G L L A L L V K D
 48050 48070 48090
 CATCGGCGGCACGTTGCACGCGCTCGTGCAGCTGCGCACCGAGGCGGGCGGGATGGACGT
 I G G T L H A L V Q L R T E A G G M D V
 48110 48130 48150
 CGCCGAGCTGGCGCCTACGGTGCAGTGCAGCCCGACAACACTACGCCGACGCGCCCGAGGA
 A E L A P T V H C Q P D N Y A D A P E E
 48170 48190 48210
 GTTCCGACCGGCCTATGTGGACTACGTGTTGAACGTGCCGCGCTCGCAGGTCCGCTACGA
 F R P A Y V D Y V L N V P R S Q V R Y D
 48230 48250 48270
 CGCATGGCACTCCGAGGAGGGCGGCCGTTCTACCGCAACGAGAACCAGGTACATGCTGAT
 A W H S E E G G R F Y R N E N R Y M L I
 48290 48310 48330
 CGAGGTGCCCCGCCGACTTCGACGCCAGTGGCGCTCCCGACCACCGGTGGATGACCTTCGA
 E V P A D F D A S A A P D H R W M T F D
 48350 48370 48390
 CCAGATCACCTACCTGCTCGGGCACAGCCACTACGTCAACATCCAGCTGCGCAGCATCAT
 Q I T Y L L G H S H Y V N I Q L R S I I
 48410 48430 48450
 CGCGTGCGCCTCGGCCGTCTACACCAGGACCGCCGGATGAAACGCGCGCTGACCGACCTG
 A C A S A V Y T R T A G *
 M K R A L T D L
 ORF17 --->
 48470 48490 48510
 GCGATCTTCGGCGGCCCGGAGGCATTCTGCACACCCTCTACGTGGGCGAGGCCGACCGTC
 A I F G G P E A F L H T L Y V G R P T V
 48530 48550 48570
 GGGGACCGGGAGCGGTTCTTCGCCCGCCTGGAGTGGGCGCTGAACAACAACCTGGCTGACC
 G D R E R F F A R L E W A L N N N W L T
 48590 48610 48630
 AACGGCGGACCACTGGTGCAGCGAGTTCGAGGGCCGGGTGCGCGACCTGGCGGGTGTCCGC
 N G G P L V R E F E G R V A D L A G V R
 48650 48670 48690
 CACTGCGTGGCCACCTGCAACGCGACGGTTCGCGCTGCAACTGGTGGTGGCGCGAGCGAC
 H C V A T C N A T V A L Q L V L R A S D
 48710 48730 48750
 GTGTCCGGCGAGGTGCTCATGCCTTCGATGACGTTTCGCGGCCACCGCGCACGCGGCGAGC
 V S G E V V M P S M T F A A T A H A A S

FIGURE 3

16/60

48770 48790 48810
 TGGCTGGGGCTGGAACCGGTGTTCTGCGACGTGGACCCCGAGACCGGCCTGCTCGACCCC
 W L G L E P V F C D V D P E T G L L D P
 48830 48850 48870
 GAGCACGTGCGCTGCTGGTGACACCGCGGACGGGCGCGATCATCGGCGTGCACCTGTGG
 E H V A S L V T P R T G A I I G V H L W
 48890 48910 48930
 GGCAGGCCCGCTCCGGTTCGAGGCGCTGGAGAAGATCGCCGCCGAGCACCAGGTCAAATC
 G R P A P V E A L E K I A A E H Q V K L
 48950 48970 48990
 TTCTTCGACGCCGCGCACGCGCTGGGCTGCACCGCCGGCGGGCGGCGGCGGCGCTTC
 F F D A A H A L G C T A G G R P V G A F
 49010 49030 49050
 GGCAACGCCGAGGTGTTTCAGCTTCCACGCCACGAAGGCGGTACCTCGTTCGAGGGCGGG
 G N A E V F S F H A T K A V T S F E G G
 49070 49090 49110
 GCCATCGTCACCGACGACGGGCTGCTGGCCGACCGCATCCGCGCCATGCACAACCTTCGGG
 A I V T D D G L L A D R I R A M H N F G
 49130 49150 49170
 ATCGCACCGGACAAGCTGGTGACCGATGTGCGGCACCAACGGCAAGATGAGCGAGTGCGCC
 I A P D K L V T D V G T N G K M S E C A
 49190 49210 49230
 GCGGCGATGGGCCTCACCTCGCTCGACGCTTTCGCCGAGACCAGGGTGCACAACCGCCTC
 A A M G L T S L D A F A E T R V H N R L
 49250 49270 49290
 AACCACGCGCTCTACTCCGACGAGCTCCGCGACGTGCGCGGCATATCCGTGCACGCGTTC
 N H A L Y S D E L R D V R G I S V H A F
 49310 49330 49350
 GATCCTGGCGAGCAGAACAACACTACCACTACGTGATCATCTCGGTGGACTCCGCGGCCACC
 D P G E Q N N Y Q Y V I I S V D S A A T
 49370 49390 49410
 GGCATCGACCGCGACCGAGTTGACGGCGATCCTGCGAGCGGAGAAGGTTGTGGCACAACCC
 G I D R D Q L Q A I L R A E K V V A Q P
 49430 49450 49470
 TACTTCTCCCCCGGGTGCCACCAGATGCAGCCGTACCGGACCGAGCCGCGCTGCGGGCTG
 Y F S P G C H Q M Q P Y R T E P P L R L
 49490 49510 49530
 GAGAACACCGAACAGCTCTCCGACCGGGTGCTCGCGCTGCCCCACCGGCCCCCGGGTGTCC
 E N T E Q L S D R V L A L P T G P A V S
 49550 49570 49590
 AGCGAGGACATCCGGCGGGTGTGCGACATCATCCGGCTCGCCGCCACCGCGGCGAGCTG
 S E D I R R V C D I I R L A A T S G E L
 49610 49630 49650
 ATCAACGCGCAATGGGACCAGAGGACGCGCAACGGTTCGTGACGACCTGCGCCACAAGTG
 I N A Q W D Q R T R N G S
 49670 49690 49710
 CCAGGAGGTTGCTCCCCGATGAACACAACCTCGTACGGCAACCGCCAGGAAGCGGGGGT
 M N T T R T A T A Q E A G V
 ORF18 --->
 49730 49750 49770
 CGCCGACGCGGCGCGCCCGGACGTGACCGGCGGGCGGTGCTGCGGGCGCTGAGCTCGGA
 A D A A R P D V D R R A V V R A L S S E

FIGURE 3

17/60

49790 49810 49830
GGTCTCCCGCGTCACCGGCGCCGGTGACGGTGACCGCGACGTGCAGGCCCGCCCGGCTCGC
V S R V T G A G D G D A D V Q A A R L A
49850 49870 49890
CGACCTCGCCGCGCACTACGGGGCGCACCCGTTACGCGCGCTGGAGCAGACGCGTGCGCG
D L A A H Y G A H P F T P L E Q T R A R
49910 49930 49950
GCTCGGCCTGGACCGCGCGGAGTTCGCCACCTGCTCGACCTGTTCCGGCCGCATCCCGGA
L G L D R A E F A H L L D L F G R I P D
49970 49990 50010
CCTGGGCACCGCGGTGGAGCACGGTCCGGCGGGCAAGTACTGGTCCAACACGATCAAGCC
L G T A V E H G P A G K Y W S N T I K P
50030 50050 50070
GCTGGACGCGCGCAGGCGCACTGGACGCGCGGTCTACCGCAAGCCTGCCTTCCCTACAG
L D A A G A L D A A V Y R K P A F P Y S
50090 50110 50130
CGTCGGCCTGTACCCCGGGCGACGTGCATGTTCCGCTGCCACTTCTGCGTGCGGGTGAC
V G L Y P G P T C M F R C H F C V R V T
50150 50170 50190
CGGTGCCCGCTACGAGGCGCATCGGTCCCGCGGGCAACGAGACGCTGGCCGCGATCAT
G A R Y E A A S V P A G N E T L A A I I
50210 50230 50250
CGACGAGGTGCCCCACGGACAACCCGAAGGCGATGTACATGTCCGGCGGGCTCGAGCCGCT
D E V P T D N P K A M Y M S G G L E P L
50270 50290 50310
GACCAACCCCGGTCTCGGCGAGCTGGTGTGCGACGCCCGCGGGCGCGGTTTCGACCTCAC
T N P G L G E L V S H A A G R G F D L T
50330 50350 50370
CGTCTACACCAACGCCTTCGCCCTCACCGAGCAGACGCTGAACCGCCAGCCCGGCCTGTG
V Y T N A F A L T E Q T L N R Q P G L W
50390 50410 50430
GGAGCTGGGCGCGATCCGCGACGTCCCTCTACGGGCTGAACAACGACGAGTACGAGACGAC
E L G A I R T S L Y G L N N D E Y E T T
50450 50470 50490
CACCGGCAAGCGCGCGCTTTCGAACGCGTCAAGAAGAACCTGCAGGGCTTCCTGCGGAT
T G K R G A F E R V K K N L Q G F L R M
50510 50530 50550
GCGCGCCGAGCGGGACGCGCCGATCCGGCTCGGCTTCAACCACATCATCTGCCGGGACG
R A E R D A P I R L G F N H I I L P G R
50570 50590 50610
GGCCGACCGGCTCACCGACCTCGTCGACTTCATCGCCGAGCTCAACGAGTCCAGCCCGCA
A D R L T D L V D F I A E L N E S S P Q
50630 50650 50670
ACGGCCGCTGGACTTCGTGACGGTGCGCGAGGACTACAGCGGCCGCGACGACGGCCGGCT
R P L D F V T V R E D Y S G R D D G R L
50690 50710 50730
GTCGGACTCCGAGCGCAACGAGCTGCGCGAGGGCCTGGTGCGGTTCTGCGACTACGCCGC
S D S E R N E L R E G L V R F V D Y A A
50750 50770 50790
CGAGCGGACCCCGGGCATGCACATCGACCTGGGCTACGCCCTGGAGAGCCTGCGGCGGGG
E R T P G M H I D L G Y A L E S L R R G

FIGURE 3

18/60

50810 50830 50850
 TGTGGACGCCGAGCTGCTGCGCATCCGCCGAGACGATGCGTCCCACCGCGCACCCCCA
 V D A E L L R I R P E T M R P T A H P Q
 50870 50890 50910
 GGTGCGGTGACAGATCGACCTGCTCGGCGACGTCTACCTCTACCGCGAGGCGGGCTTCCC
 V A V Q I D L L G D V Y L Y R E A G F P
 50930 50950 50970
 GGAGCTGGAGGGCGCCACCCGCTACATCGCGGGCGGGTCACCCCGTCGACCAGCCTGCG
 E L E G A T R Y I A G R V T P S T S L R
 50990 51010 51030
 CGAGGTGGTGGAGAACTTCGTGCTGGAGAACGAGGGCGTGCAGCCCCGCCCGGCGACGA
 E V V E N F V L E N E G V Q P R P G D E
 51050 51070 51090
 GTACTTCCTCGACGGCTTCGACCAGTCGGTGACCGCACGGCTCAACCAGCTCGAACGAGA
 Y F L D G F D Q S V T A R L N Q L E R D
 51110 51130 51150
 CATCGCCGACGGGTGGGAGGACCACCGCGGCTTCCTGCGCGGAAGGTGAACCGGAGTTGC
 I A D G W E D H R G F L R G R *
 51170 51190 51210
 GAGTACGTGAGCTGGCGGTGGCGGGCGGTTTCGAGTTCACCCCGACCCGAAGCAGGACC
 V A G G F E F T P D P K Q D R
 ORF19 --->
 51230 51250 51270
 GGCGGGGCCTGTTCGTGTCTCCGCTGCAGGACGAGGCGTTCGTGGGCGCGGTGGGCCATC
 R G L F V S P L Q D E A F V G A V G H R
 51290 51310 51330
 GGTTCCTCCGTCGCCCAGATGAACCACATCGTCTCCGCCCCGGGCGTGTGCGCGGGCTGC
 F P V A Q M N H I V S A R G V L R G L H
 51350 51370 51390
 ACTTCACCACCACCCCGCCGGGGCAGTGCAAGTACGTCTACTGCGCGCGCGGCCGGGCGC
 F T T T P P G Q C K Y V Y C A R G R A L
 51410 51430 51450
 TCGACGTTCATCGTCGACATCCGGGTGCGCTCGCCGACGTTCCGGGAAGTGGGACCGGTGG
 D V I V D I R V G S P T F G K W D A V E
 51470 51490 51510
 AGATGGACACCGAGCACTTCGCGGCGGTCTACTTCCCCAGGGGCACCGCGCACGCCTTCC
 M D T E H F R A V Y F P R G T A H A F L
 51530 51550 51570
 TCGCGCTTGAGGACGACACCCTGATGTGCTACCTGGTCAGCACGCCGTACGTGGCCGAGT
 A L E D D T L M S Y L V S T P Y V A E Y
 51590 51610 51630
 ACGAGCAGGCGATCGACCCGTTTCGACCCCGCGCTGGGTCTGCCGTGGCCCGCGGACCTGG
 E Q A I D P F D P A L G L P W P A D L E
 51650 51670 51690
 AGGTGCTGCTCTCCGACCGCGACACCGTGGCCGTGGACCTGGAGACCGCCAGGCGGCGAG
 V V L S D R D T V A V D L E T A R R R G
 51710 51730 51750
 GGATGCTGCCCCGACTACGCCGACTGCCTCGGCGAGGAGCCCCGCCAGCACCGGCAGGTGAC
 M L P D Y A D C L G E E P A S T G R *
 * A S Q R P S S G A L V P L H
 <--- ORF20...

FIGURE 3

00463705 508220

19/60

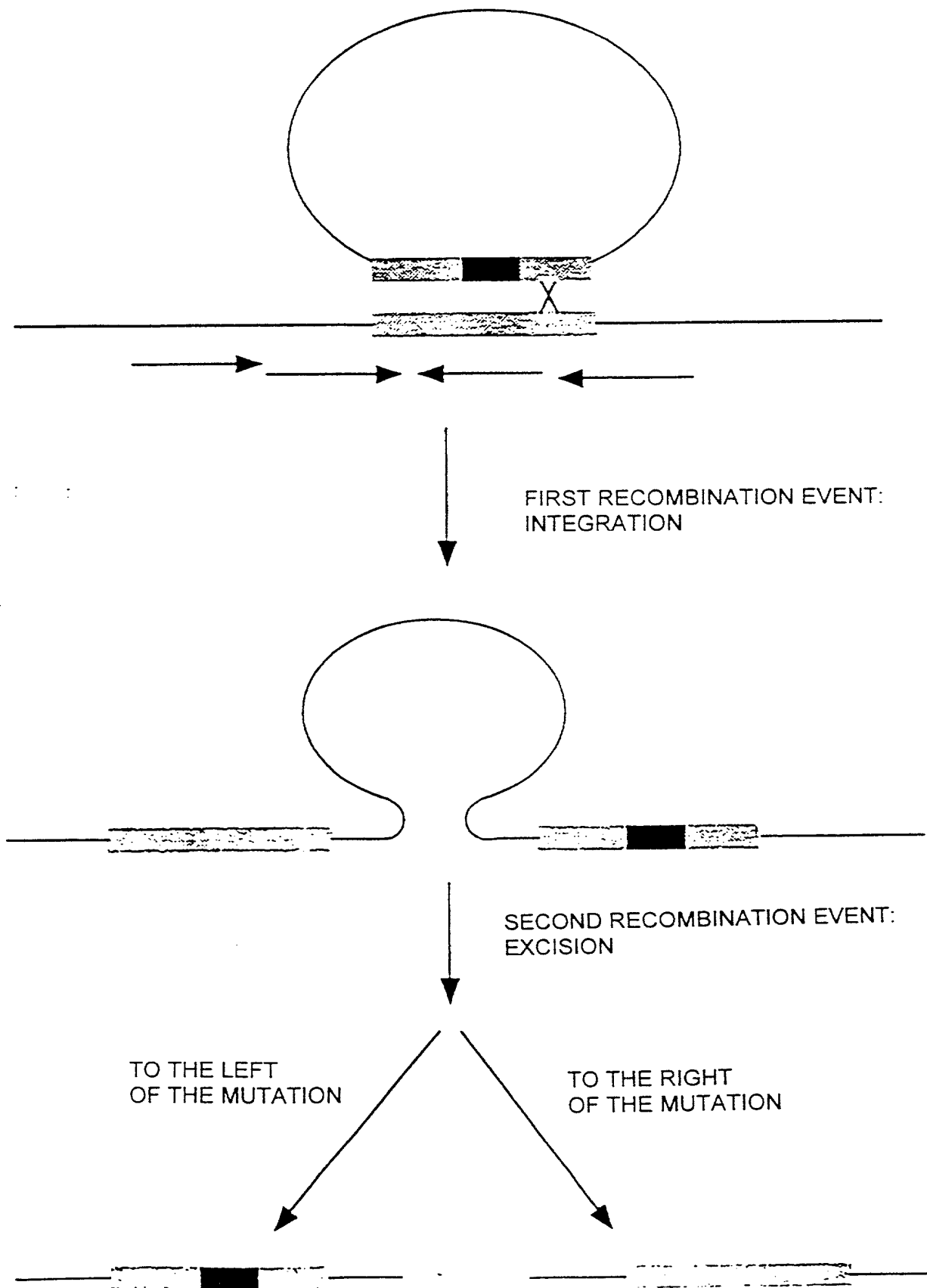


FIGURE 4

20/60

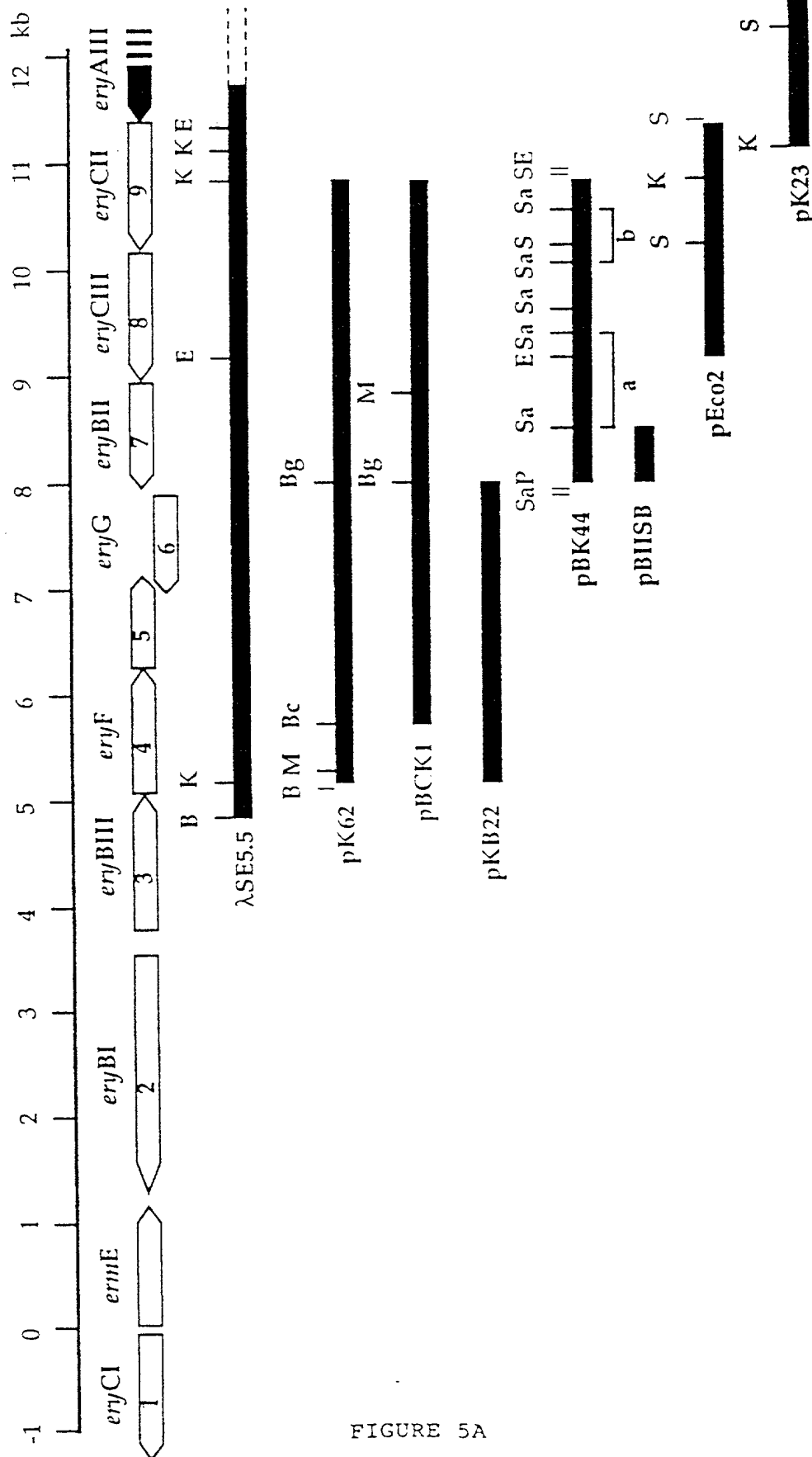


FIGURE 5A

21/60

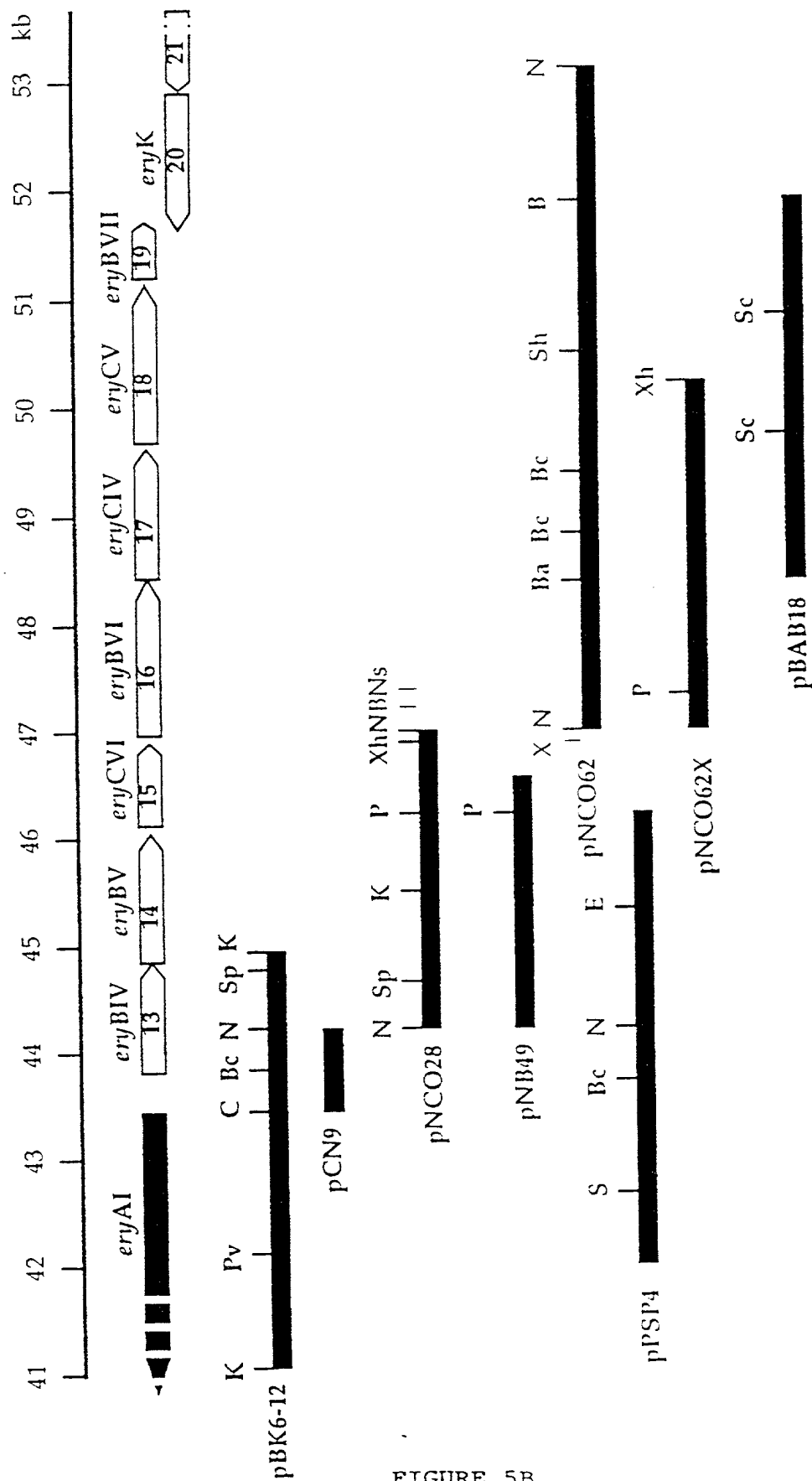


FIGURE 5B

22/60

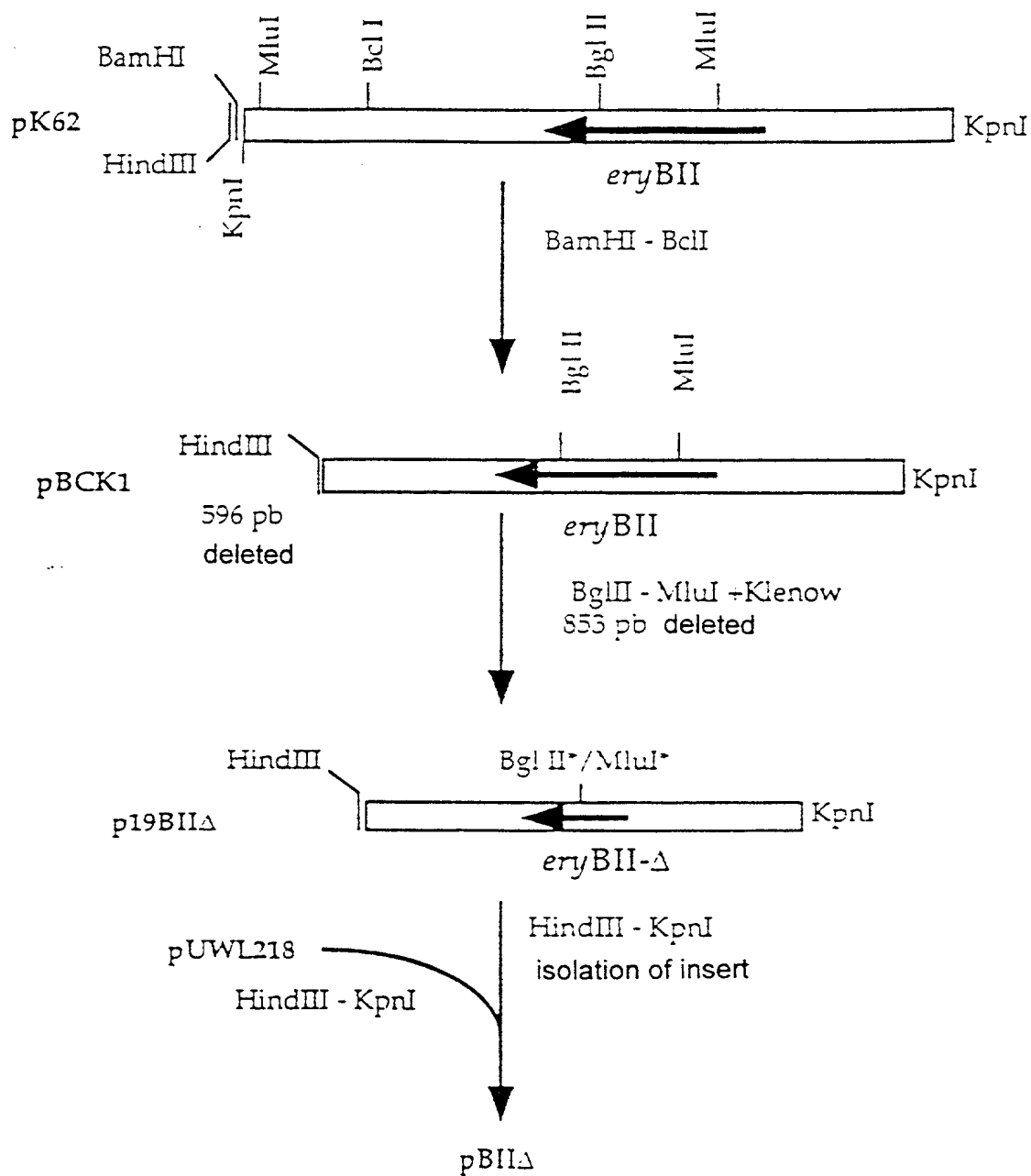


FIGURE 6A

23/60

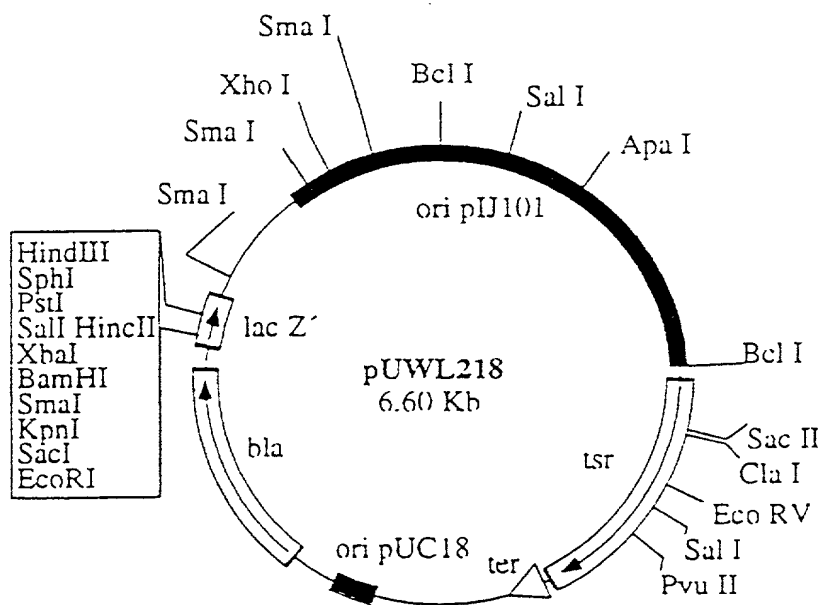


FIGURE 6B

24/60

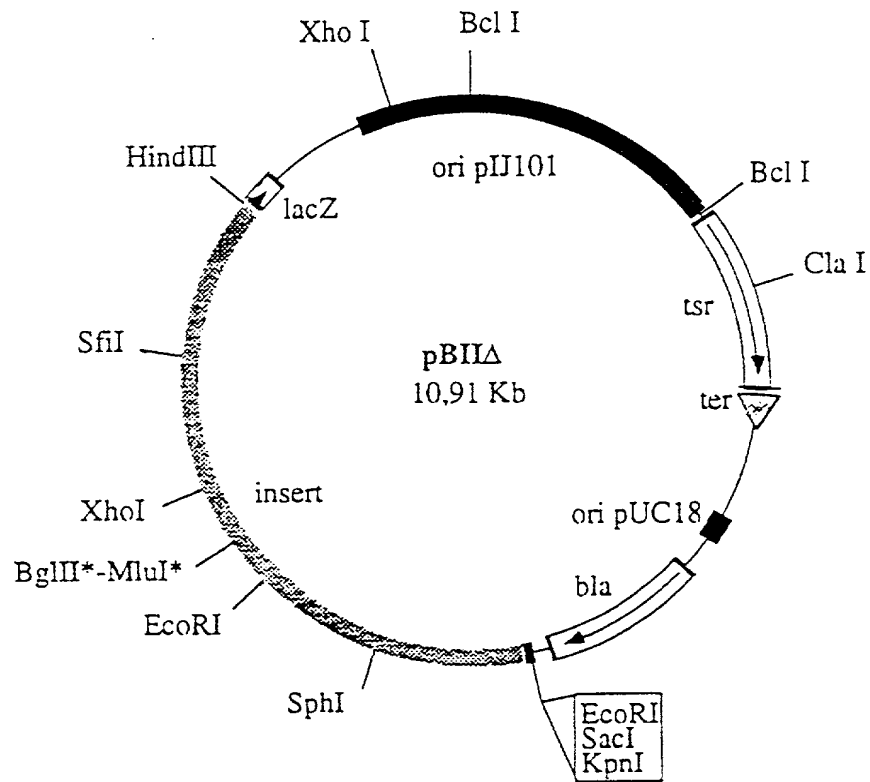


FIGURE 6C

25/60

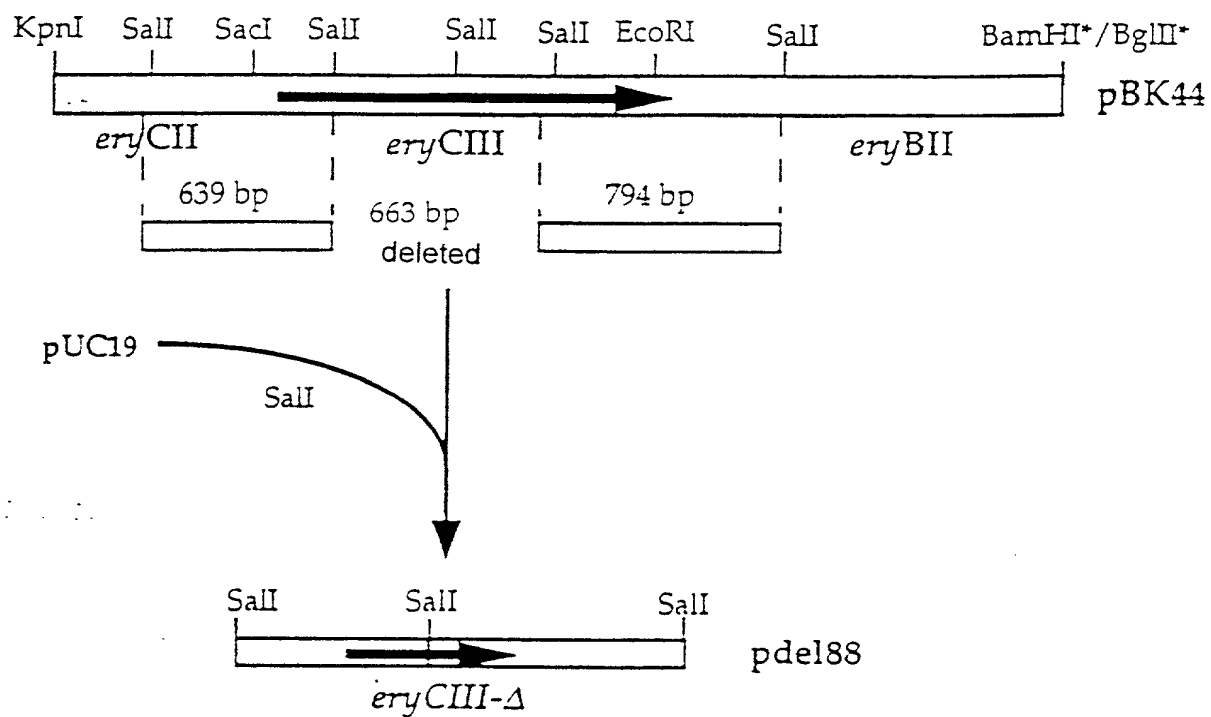


FIGURE 7A

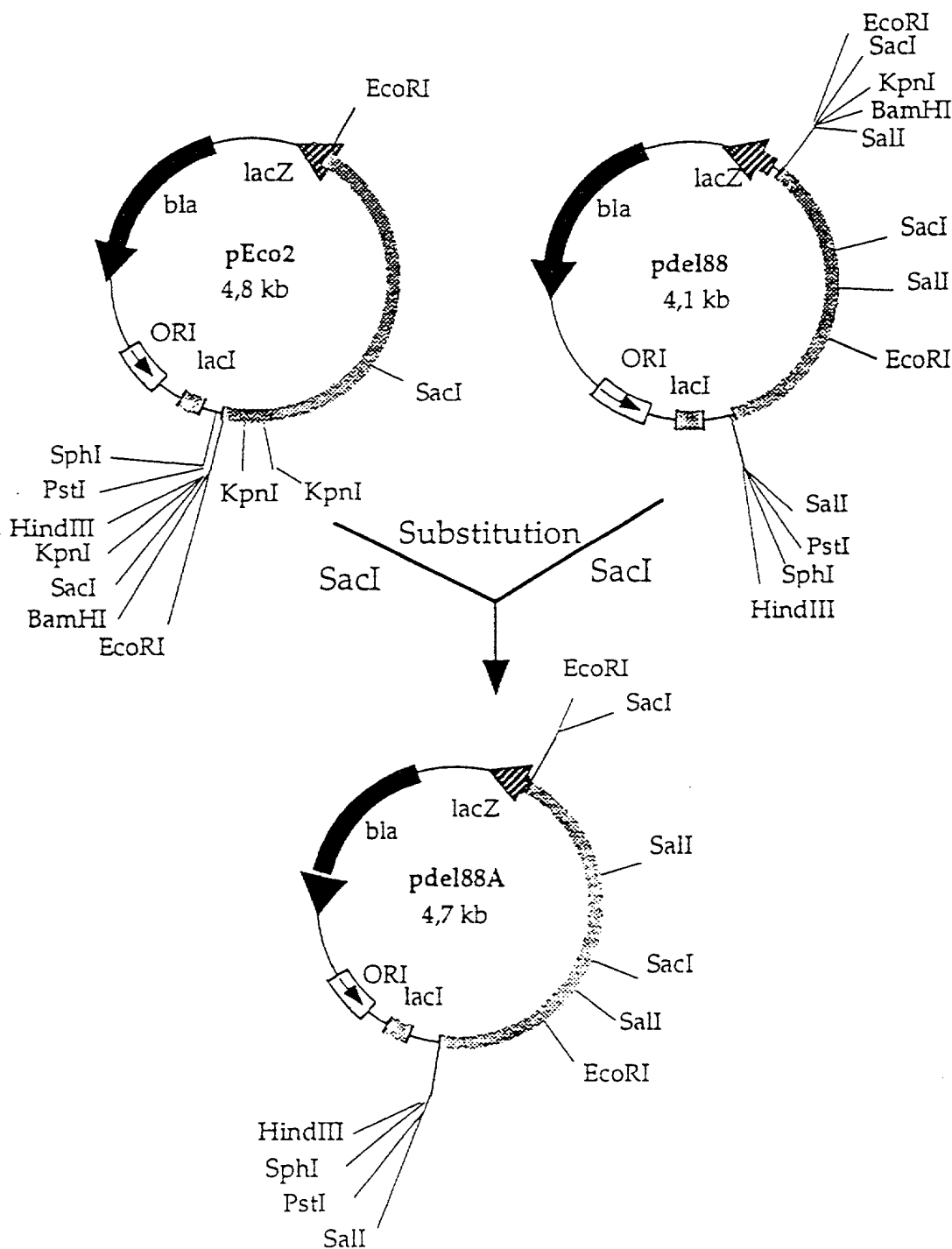


FIGURE 7B

09463705.022300



28/60

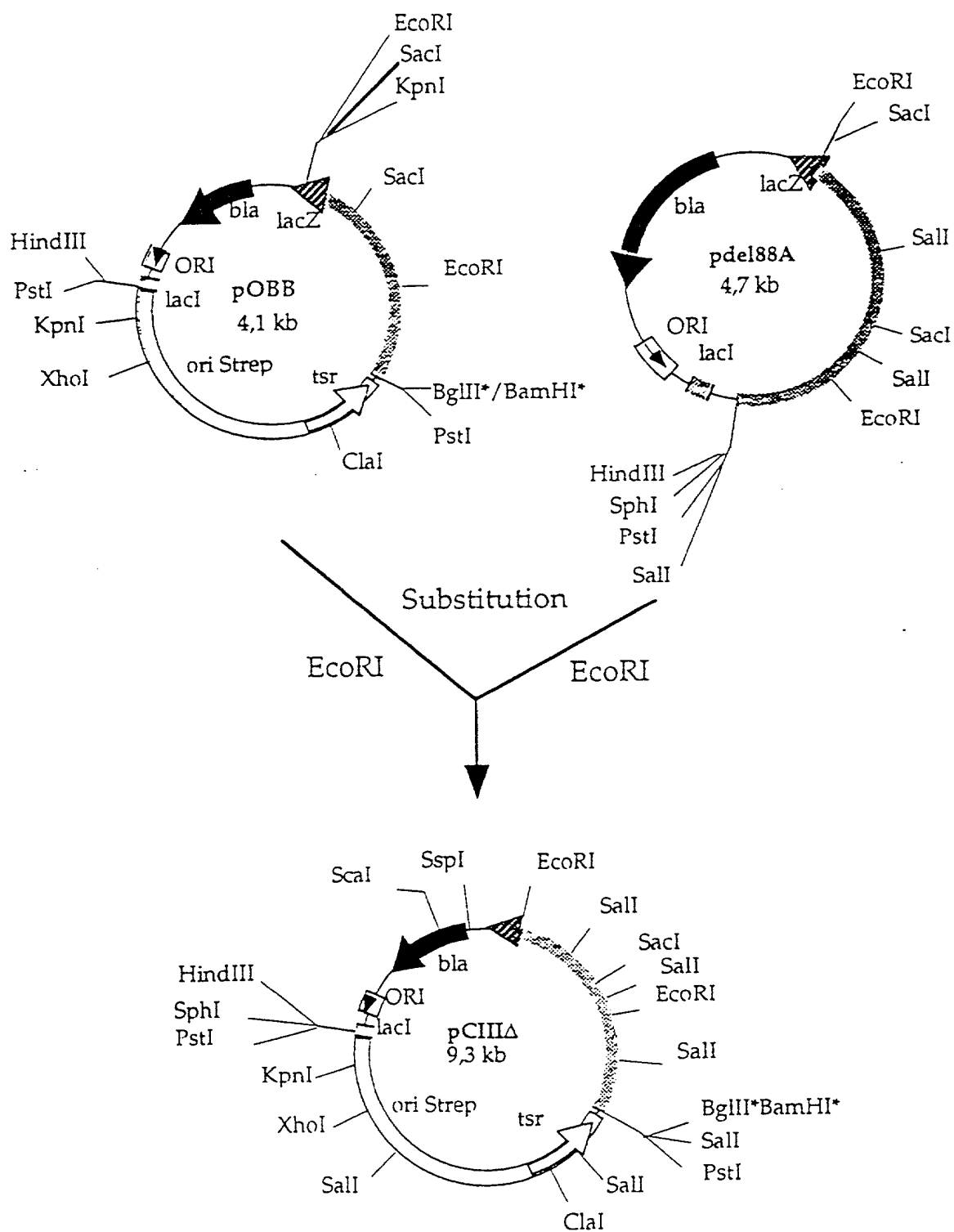


FIGURE 7D

29/60

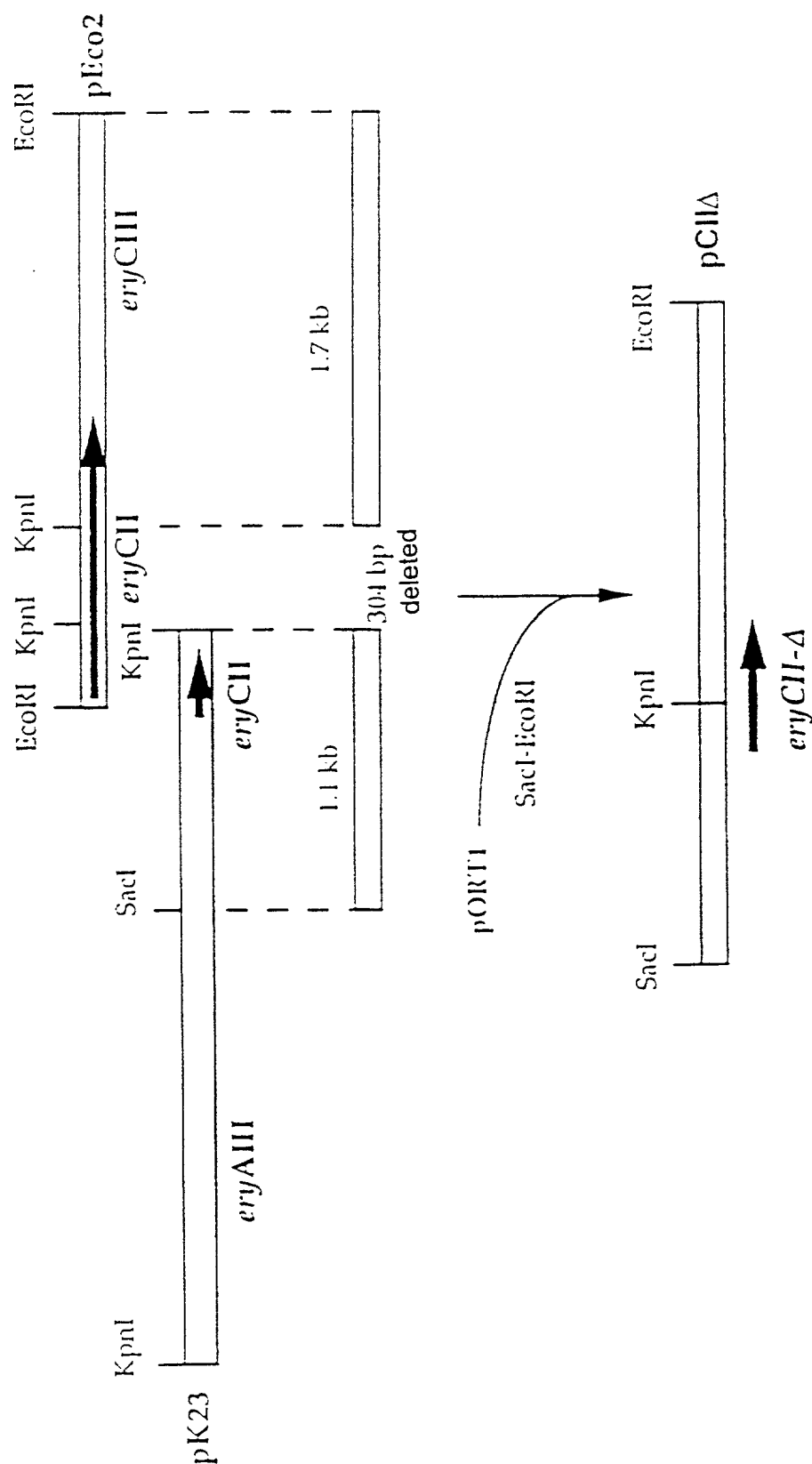


FIGURE 8A

30/60

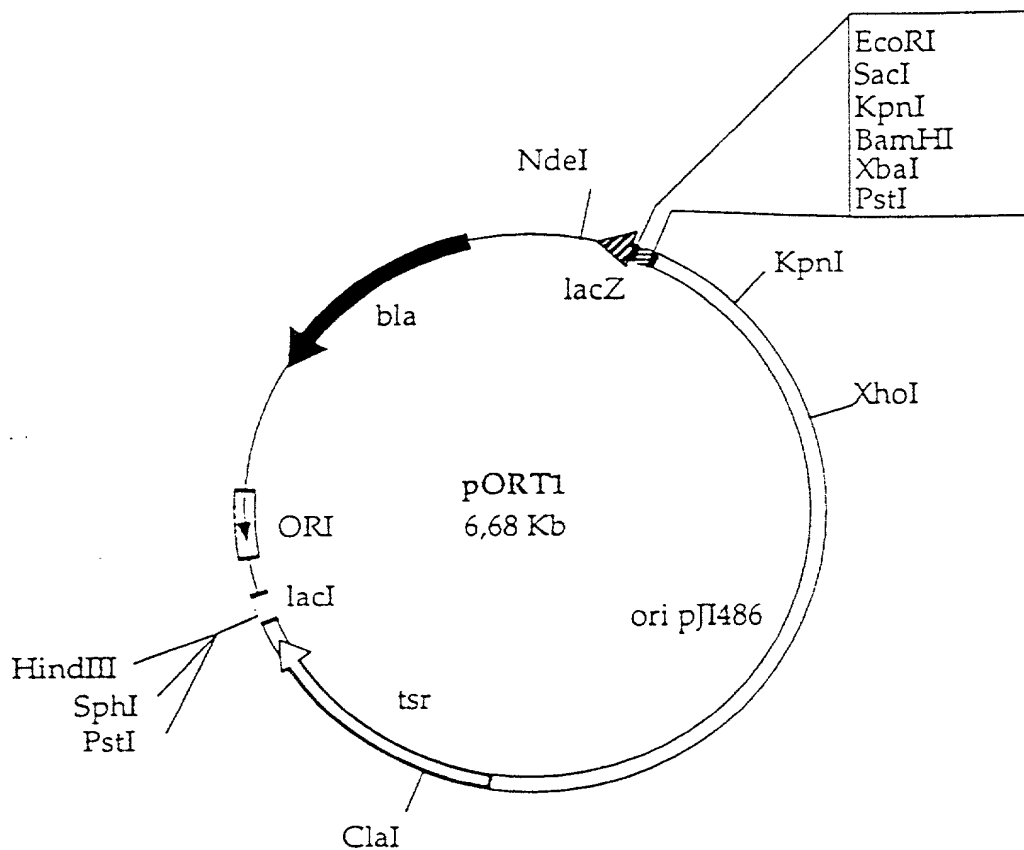


FIGURE 8B

31/60

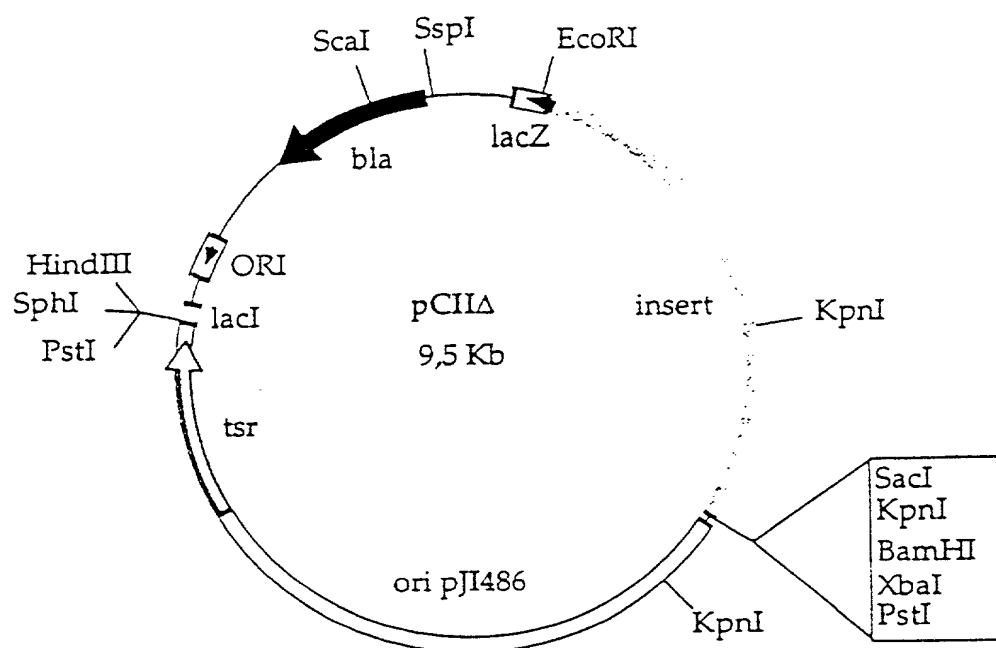


FIGURE 8C

32/60

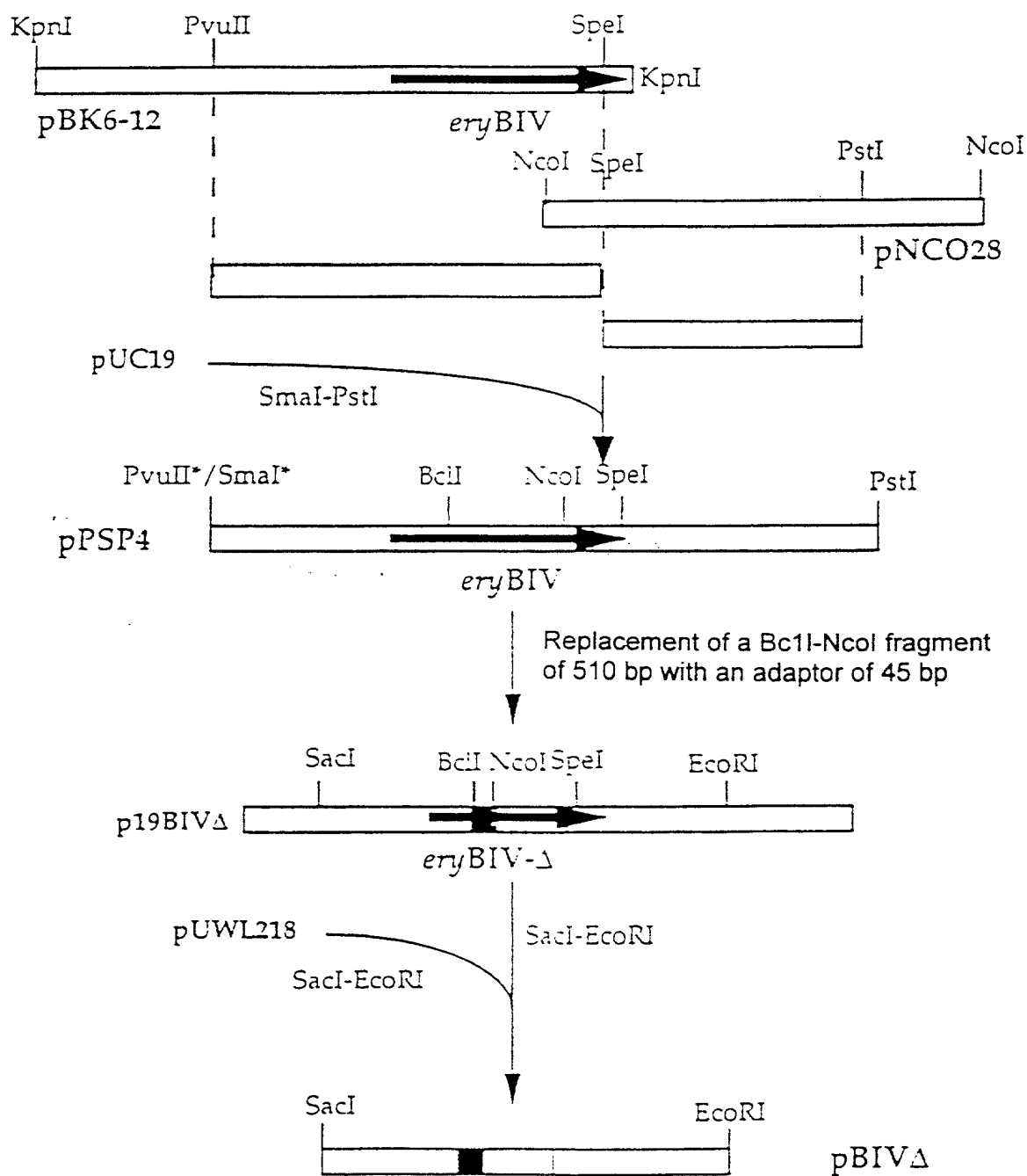


FIGURE 9A

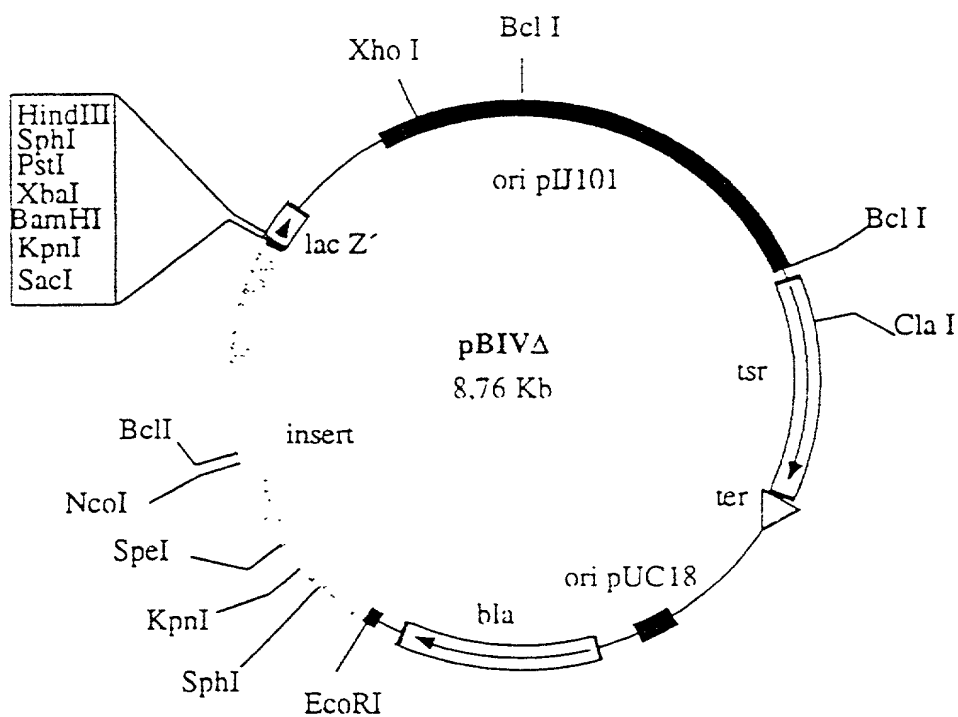


FIGURE 9B

09463705 50463705

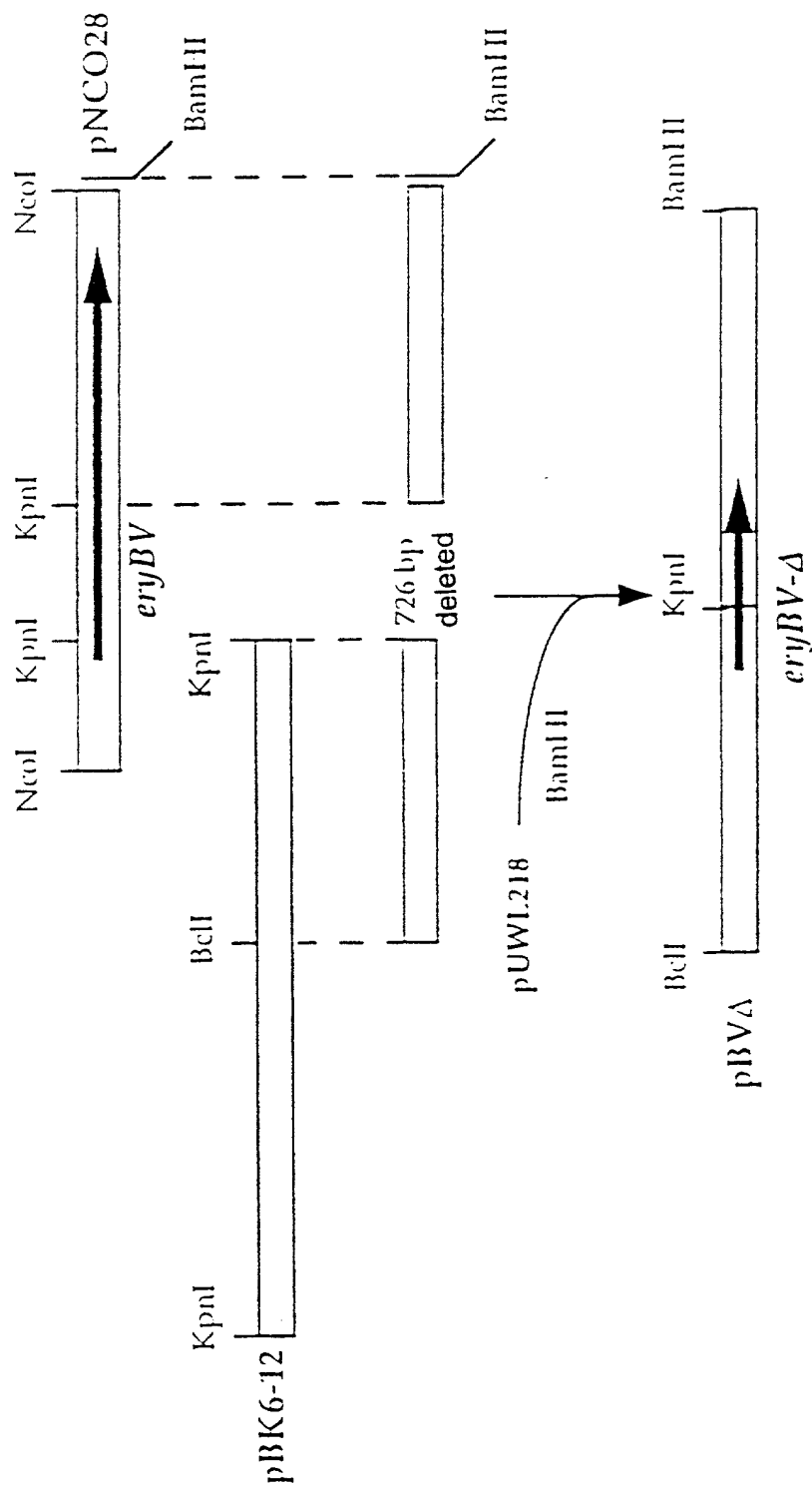


FIGURE 10A

35/60

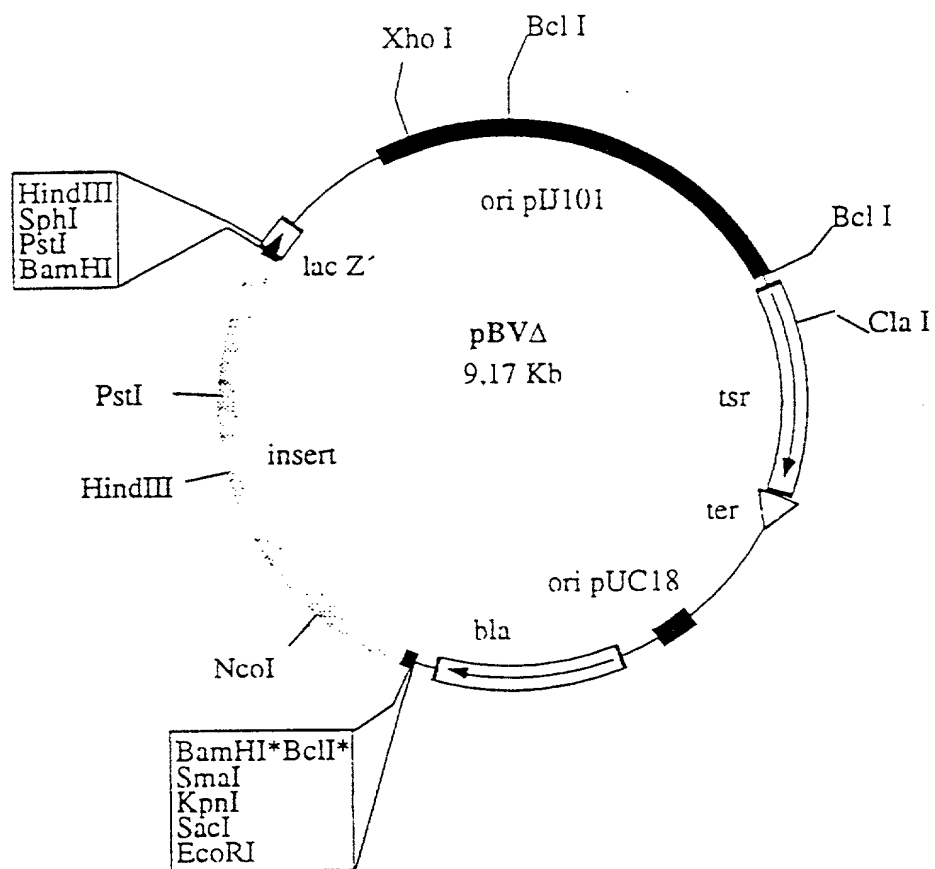


FIGURE 10B

36/60

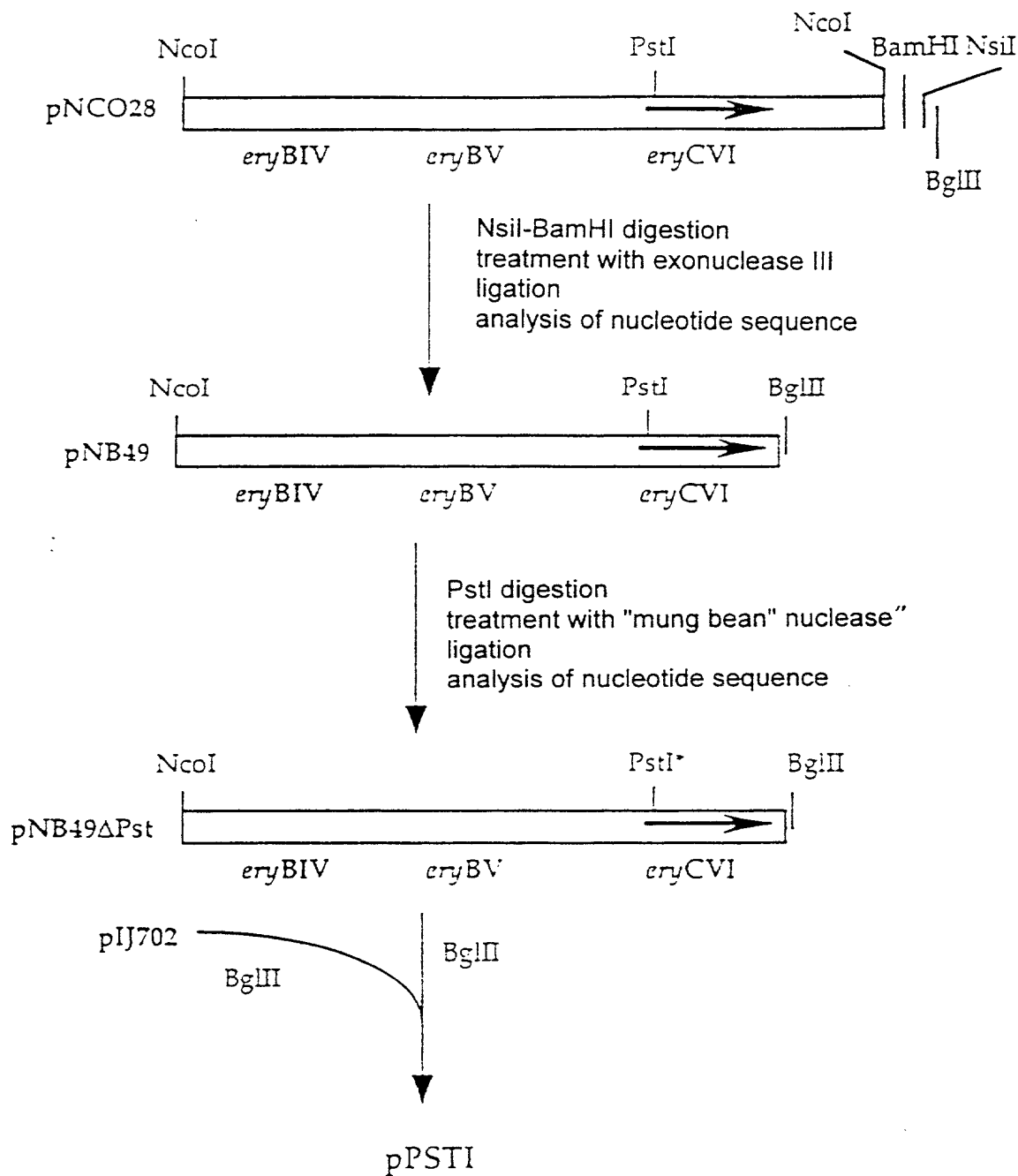


FIGURE 11A

37/60

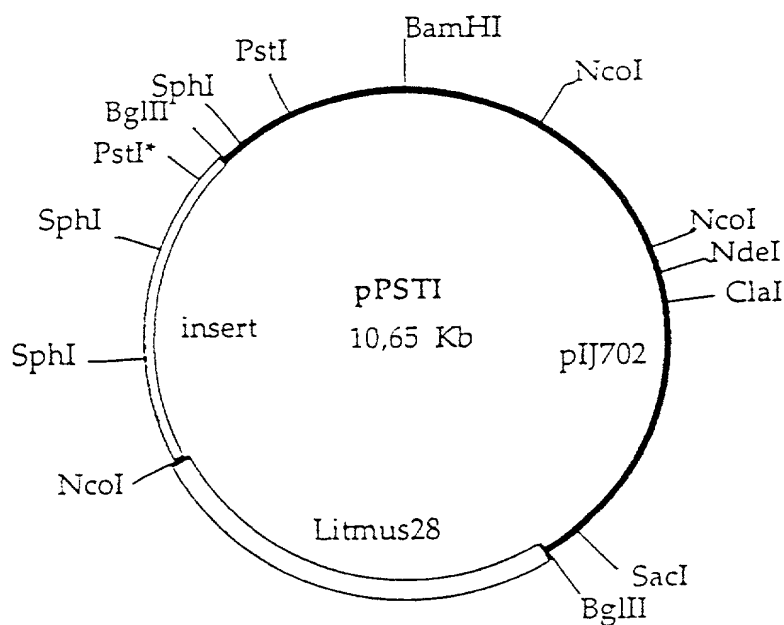


FIGURE 11B

00463705 0000

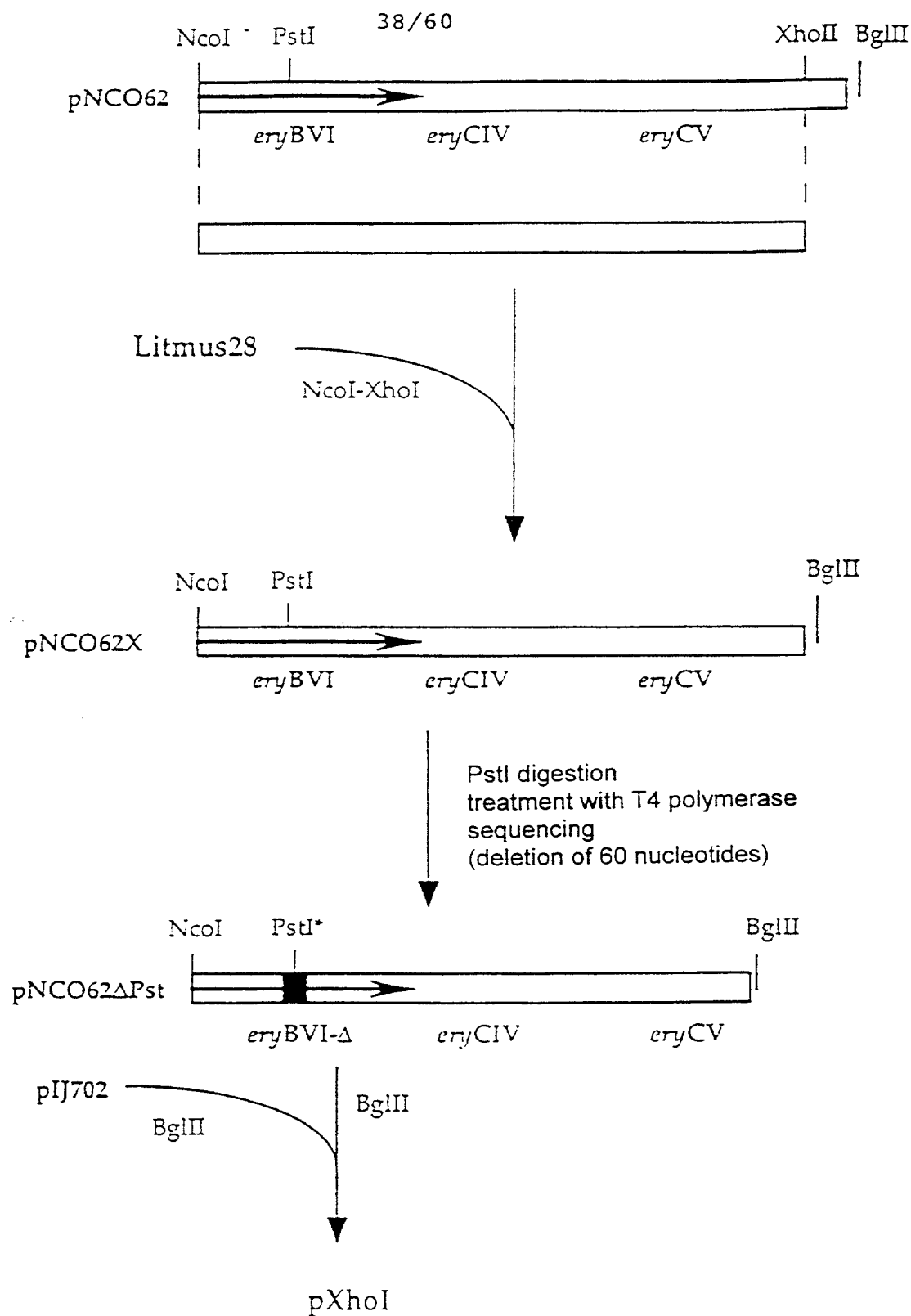


FIGURE 12A

39/60

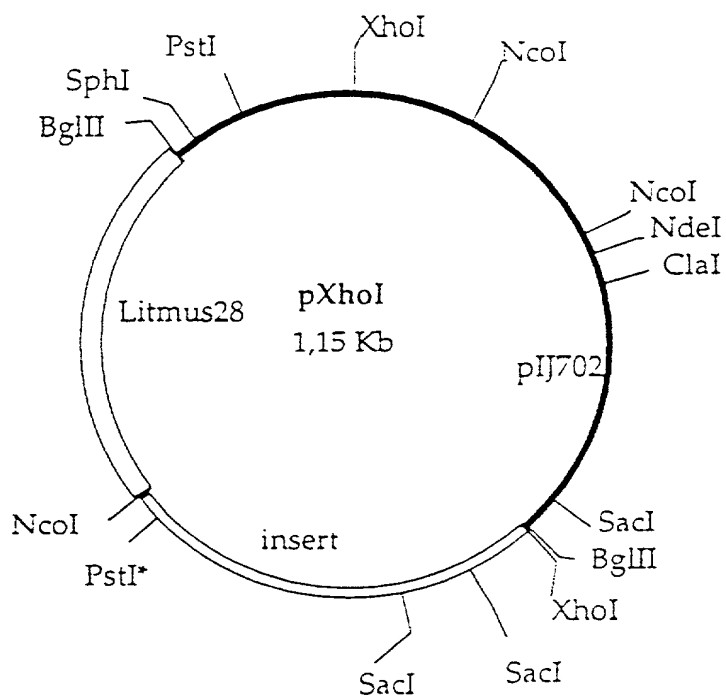


FIGURE 12B

40/60

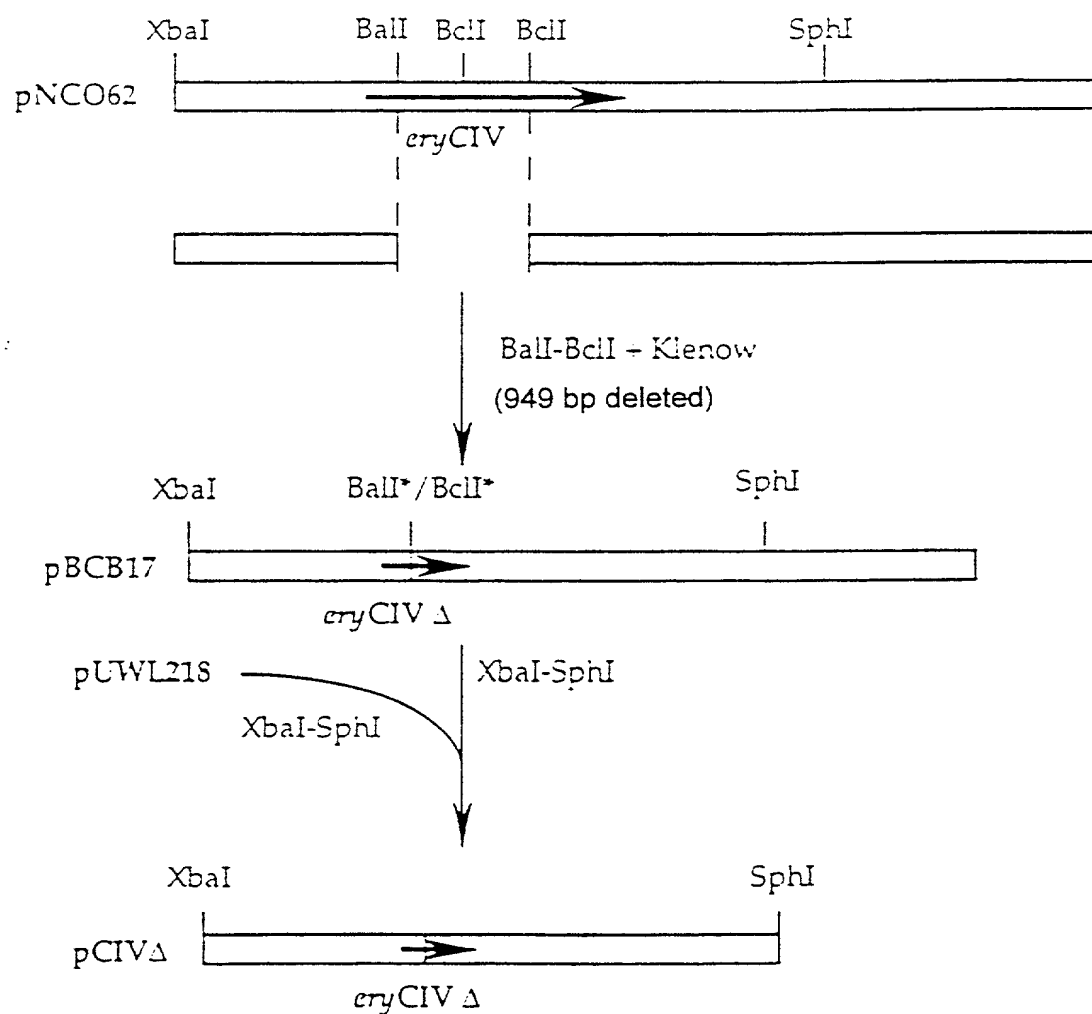


FIGURE 13A

41/60

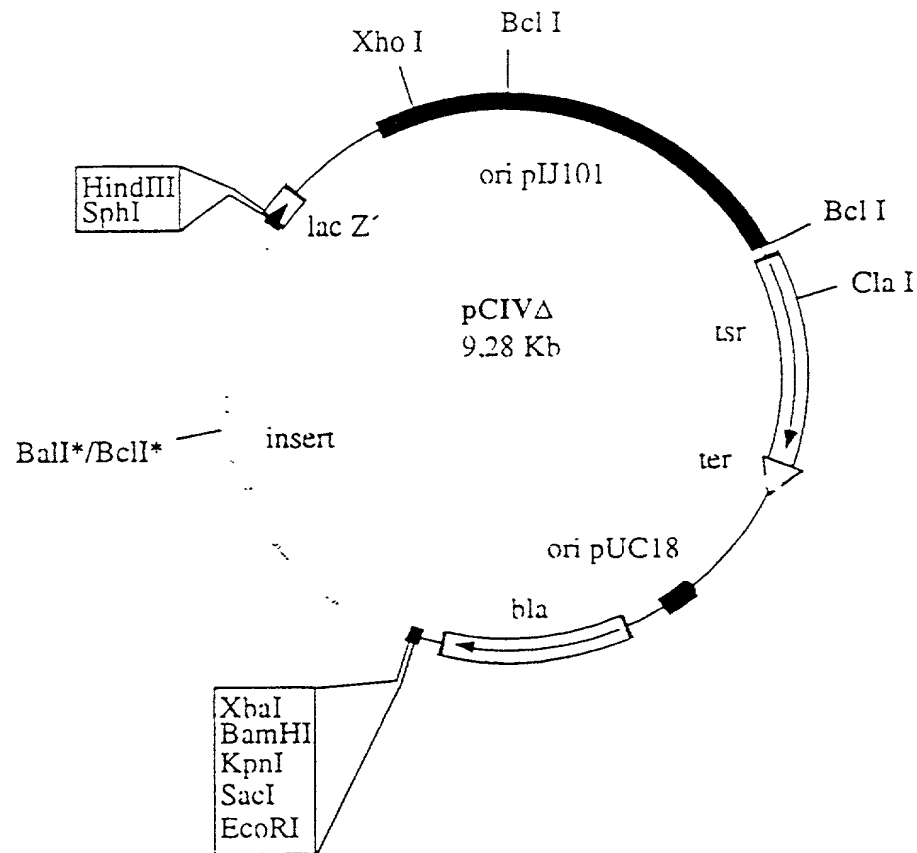


FIGURE 13B

42/60

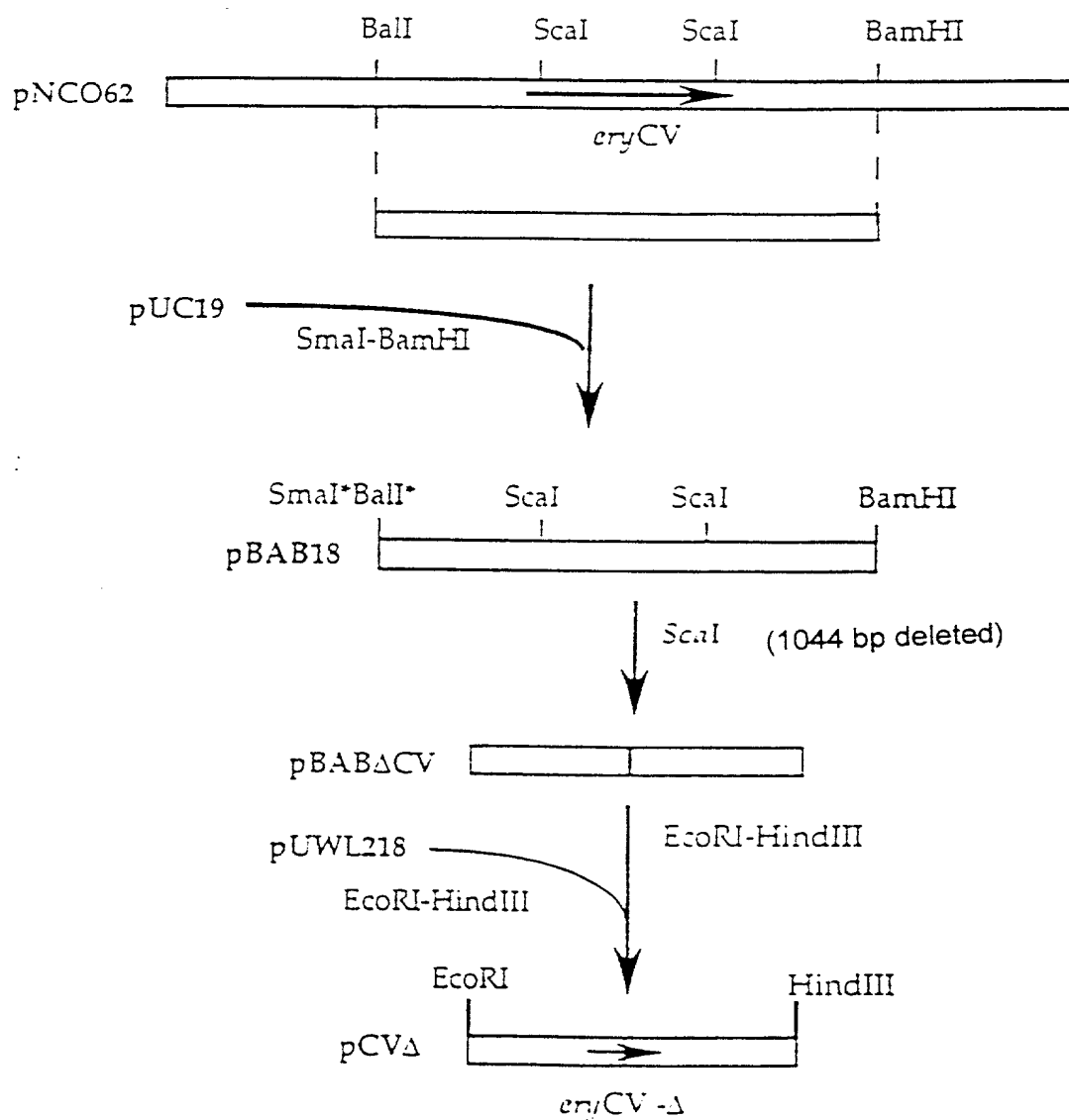


FIGURE 14A

43/60

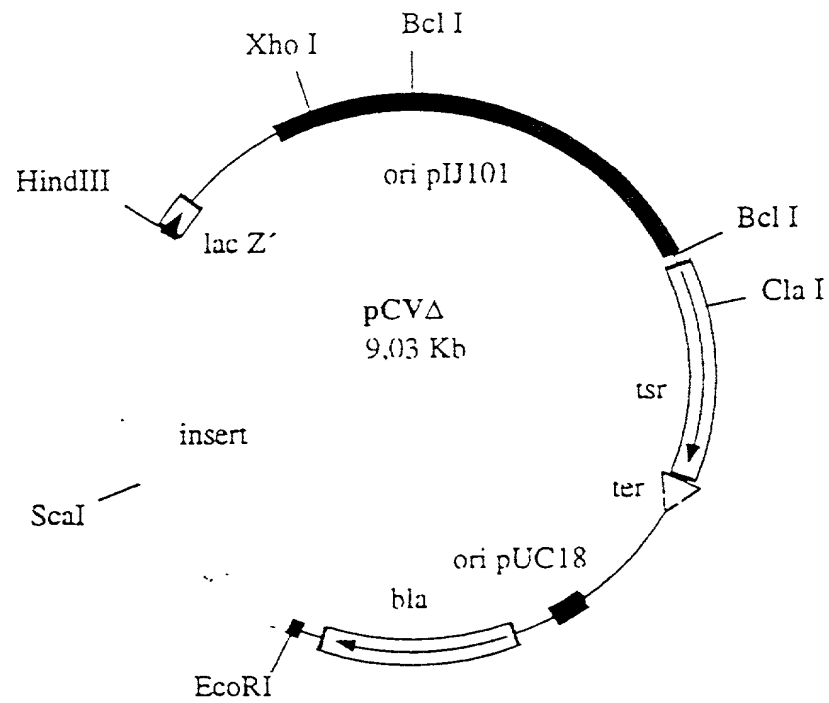


FIGURE 14B

44/60

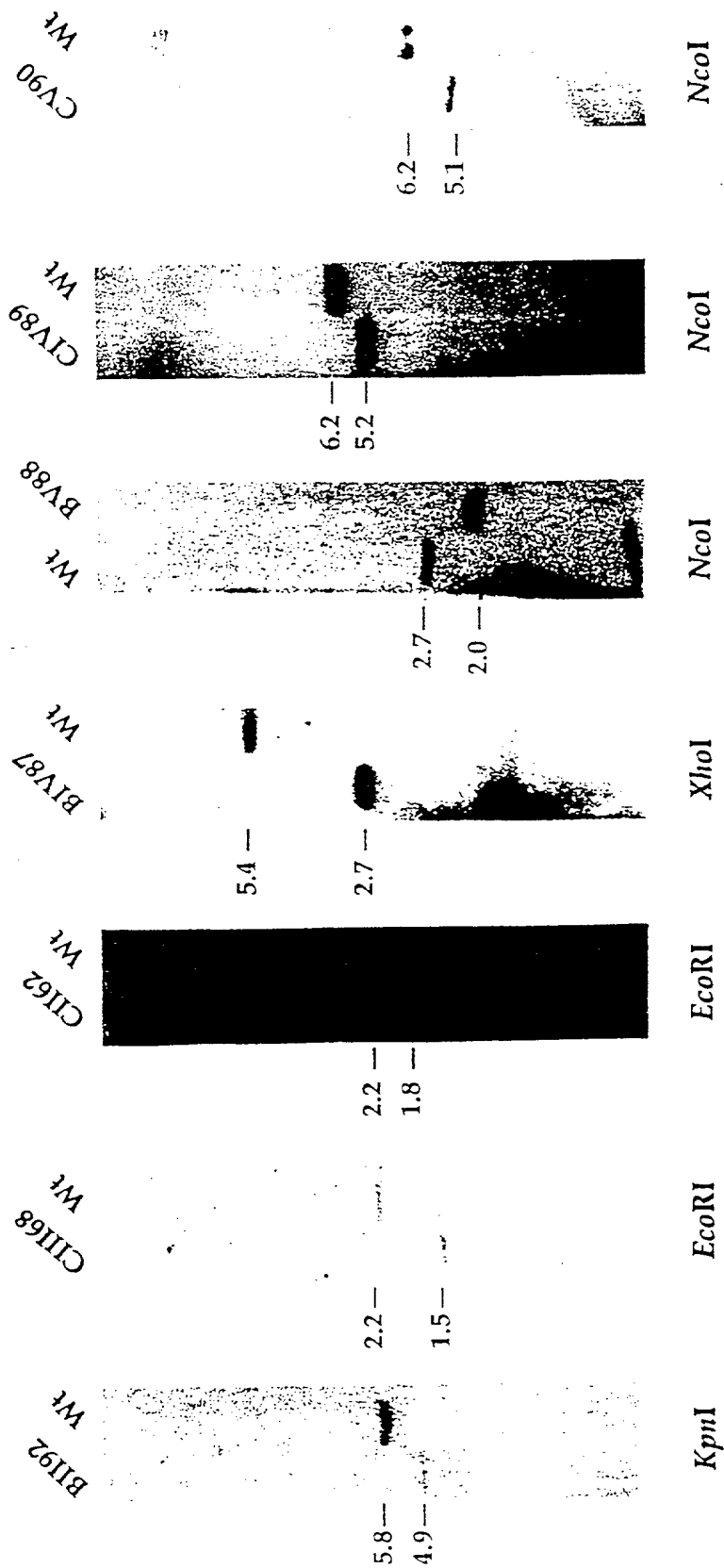


FIGURE 15

45/60

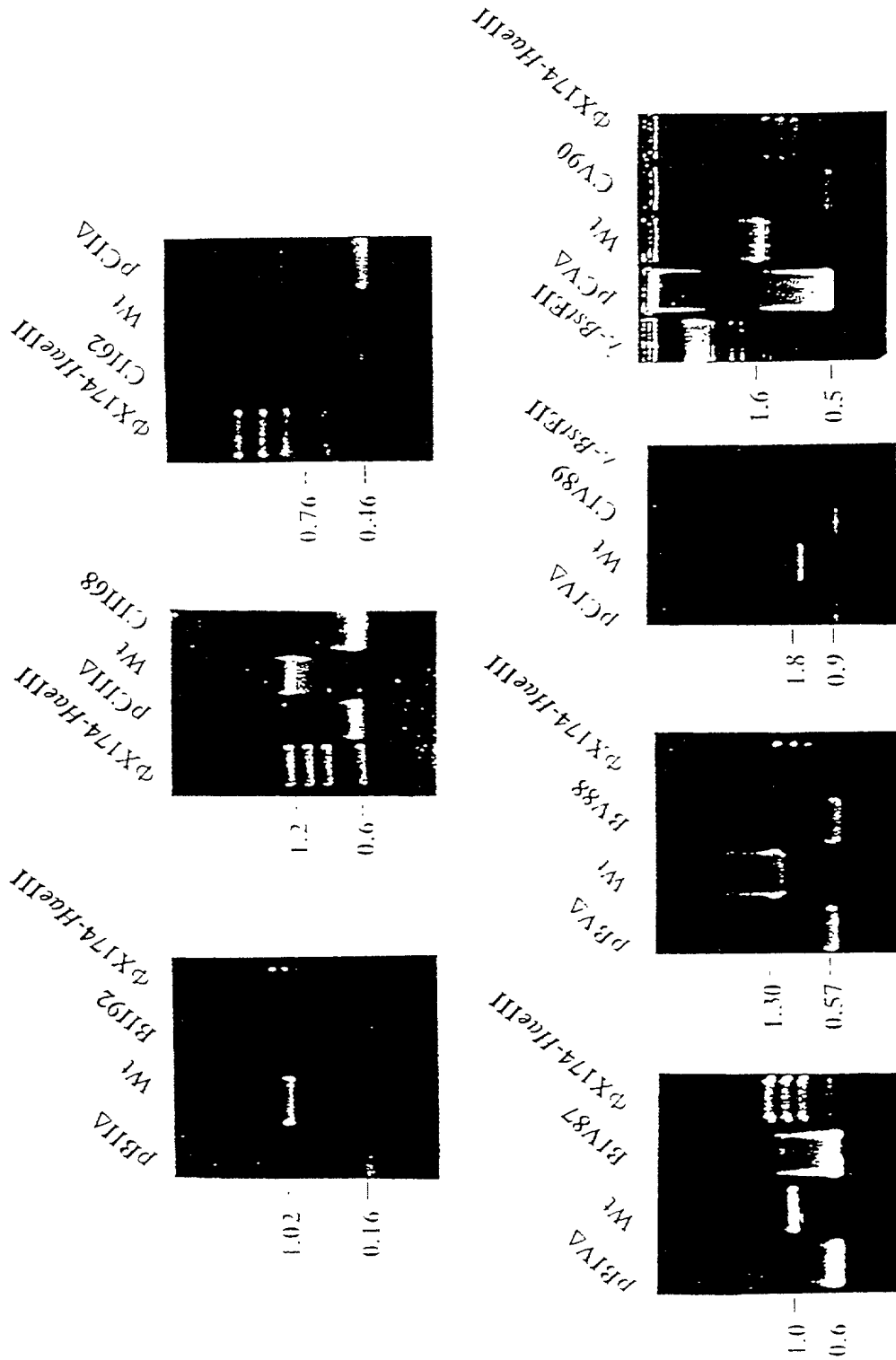


FIGURE 16

09463705-022300

46/60

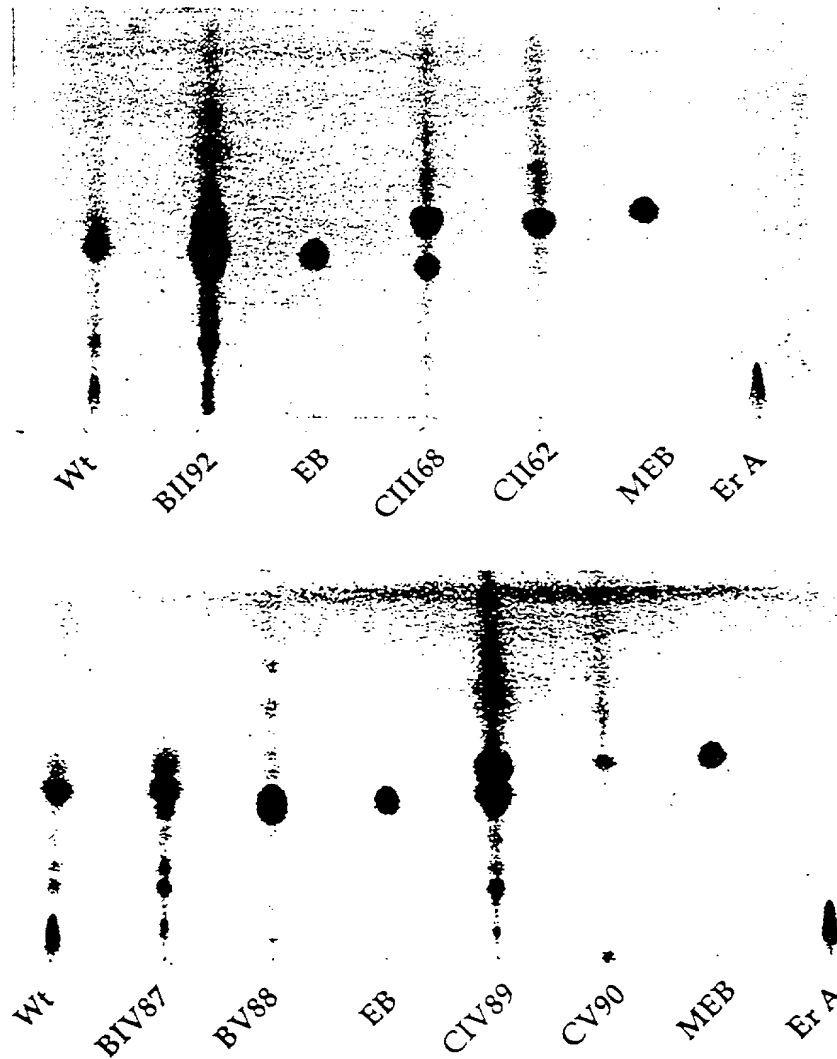


FIGURE 17

47/60

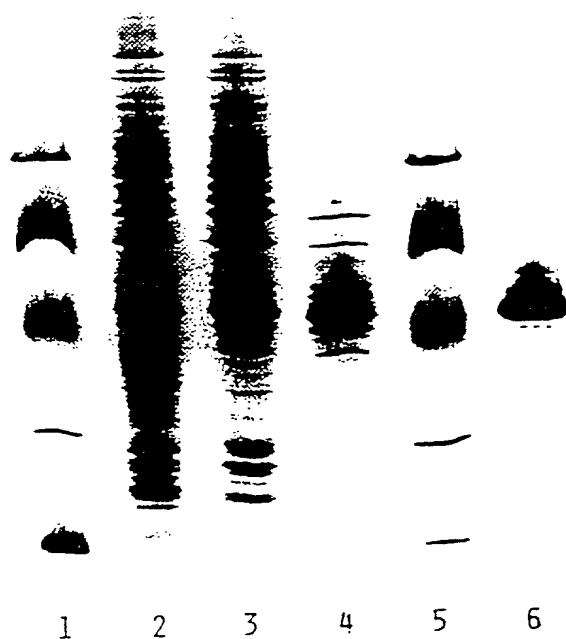


FIGURE 18

48/60

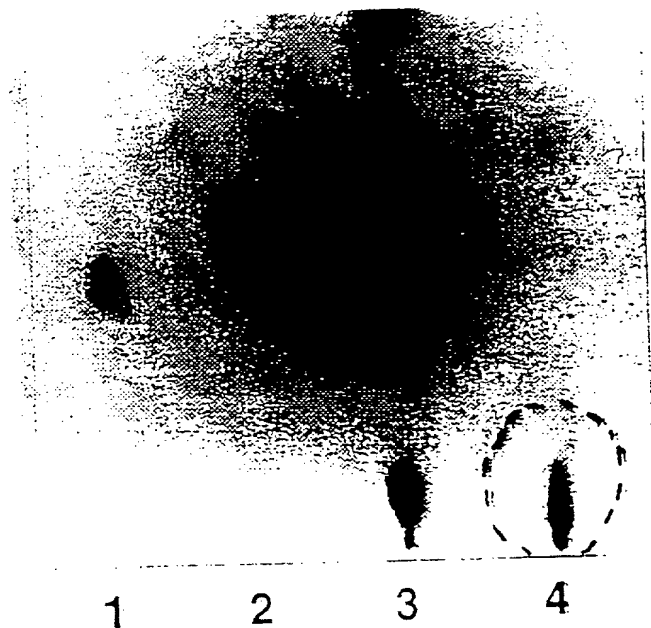


FIGURE 13

09463705.022300

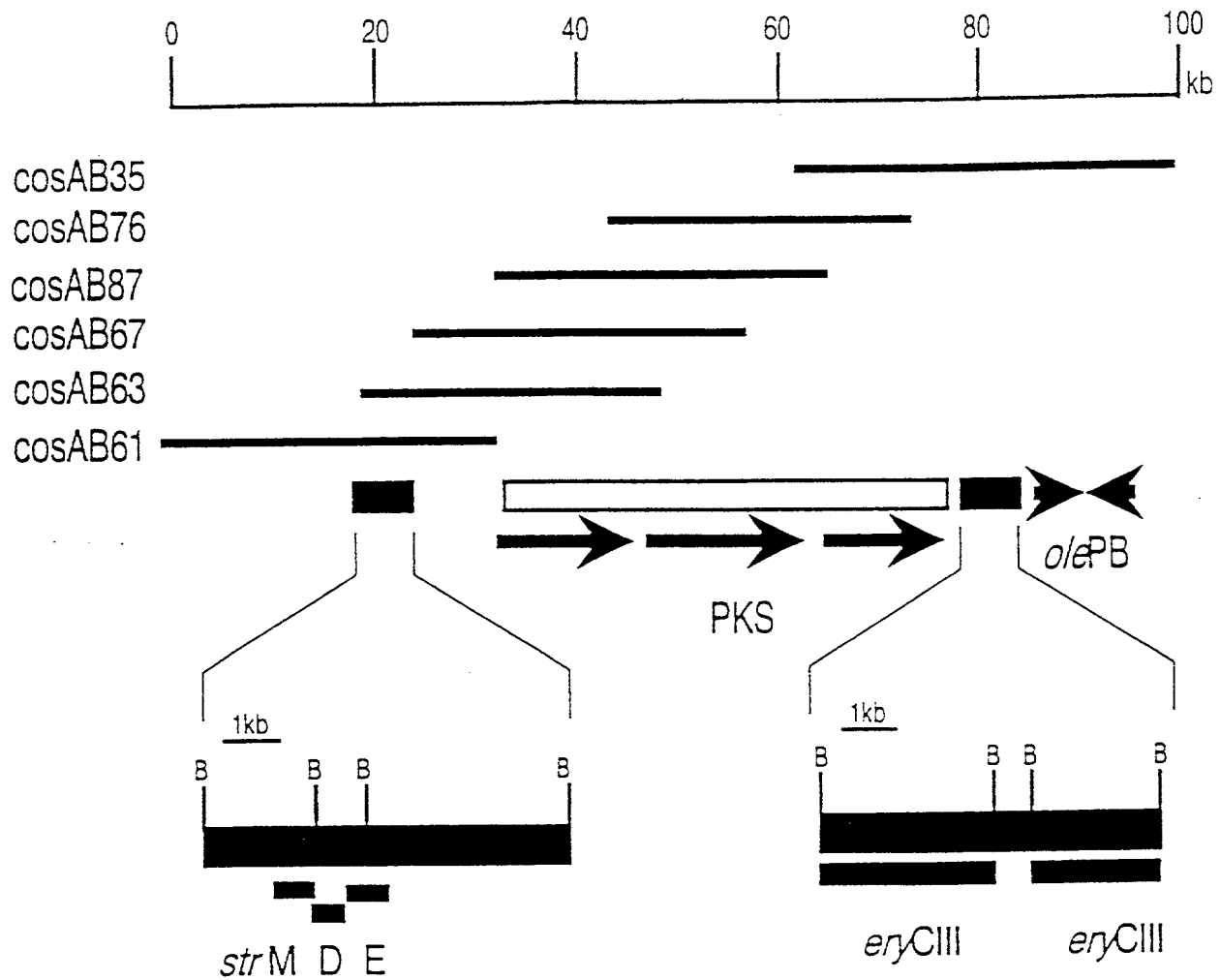


FIGURE 20

50/60

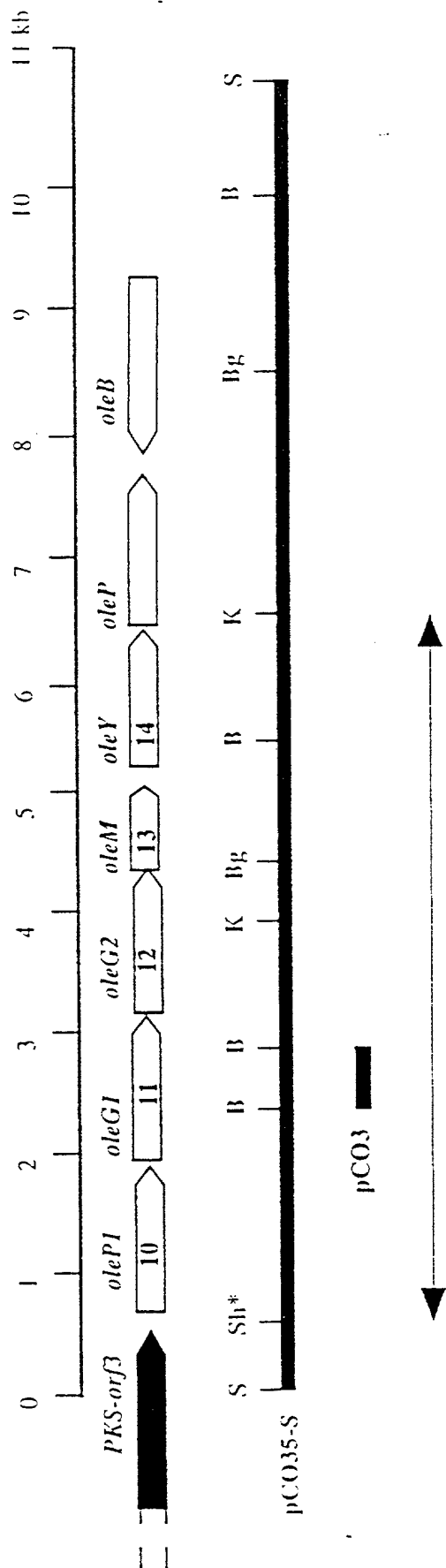


FIGURE 21

1 GCATGCCCCCGCTTTCCTCCCCCTCTCCGAACGCATCGACGACCCGATCCCCCTCAGGGAC 60
 61 CGGTGAAGGAGCGTGTTCACCTCATGTCAGGACATGCAAGGCGTACAGCCCGAACCAGCCA 120
 121 GTGTGAAACACGCGGCGGACGCAGCTCGAACAGAGCGGACGCGCACGGAAGCCGCCACG 180

*Pst*I
 181 GAGATGGAGGACAGCGAACTGGGGCGCGCCTGCAGATGCTCCGCGGCATGCAGTGGGTC 240
 M E D S E L G R R L Q M L R G M Q W V
 *Ole*P1
 241 TTCGGCGCCCAACGGCGATCCGTACGCCCGGCTGCTGTGTGGCATGGAGGATGACCCGTCA 300
 F G A N G D P Y A R L L C G M E D D P S
 301 CCTTCTACGACGCGATACGGACCCCTGGCGGAGCTGCACCGGAGCAGGACCGGAGCCTGG 360
 P F Y D A I R T L G E L H R S R T G A W
 361 GTCACCGCCGACCCCGGCTCGGGGGCGGCATCCTCGCCGACCGGAAGGCTCGGTGCCCG 420
 V T A D P G L G G R I L A D R K A R C P
 421 GAAGGCTCGTGGCCCGGTGGCGGCGAAGACCGACGGGCTGGAGCAGTACGTGCTGCCCGGG 480
 E G S W P V R A K T D G L E Q Y V L P G
 481 CACCAGGCGTTCCCTGCGGCTGGAGCGCGAGGAGCGCGGAGCGACTGCGGGAGGTCGCGGCG 540
 H Q A F L R L E R E E A E R L R E V A A
 541 CCGGTGCTGGGGCGCGGCGGTGACCGGTGGCGCGCGCGCTGATCGACGAGGTCCTGCGCG 600
 P V L G A A A V D A W R P L I D E V C A
 601 GGGCTCGGGAAGGGGCTGCCGGACACCGTTCGACCTGGTCGAGGAGTACGCGGGGCTGGTG 660
 G L A K G L P D T F D L V E E Y A G L V
 661 CCGGTGAGGTGCTGGCGCGCGGATCTGGGGCGTCCCGGAGGAGGACCGCGCGGTCGGG 720
 P V E V L A R I W G V P E E D R A R F G

FIGURE 22

721 CGTGACTGCCGGGGCGCTCGCTCCCGCGCTGGACAGCCTCCTGTGTCCCCAGCAGTTGGCG 780
R D C R A L A P A L D S L L C P Q Q L A

781 CTGAGCAAGGACATGGCGTCCGCCCTGGAGGACCTGCGTCTCTCTTCGACGGCCTCGAC 840
L S K D M A S A L E D L R L L F D G L D

841 GCGACGCGCGCCTCGCCGGGCCCGCCGACGGTGACGGAACGGCCGTGGCCCATGCTCACC 900
A T P R L A G P A D G D G T A V A M L T

901 GTTCTGCTCTGACGGAGCCGGTGACCAACGGCGGATCGGGAAACACCGTGCTCGGGCTCCTT 960
V L L C T E P V T T A I G N T V L G L L

961 CCCGGGCGAGTGGCCCGTGCCCTGCACCGCGGGGTGGCTGCCGGGCGAGTTGCCGGGCAG 1020
P G Q W P P C T G R V A A G Q V A G Q

1021 GCGCTGCACCGGGCGGTGCTGTAACCGTATCGCGACGCGGTTCCGCCGGGAGGACCTGGAG 1080
A L H R A V S Y R I A T R F A R E D L E

1081 TTGGGGGGCTGCGAGGTCAAGTCCGGTGACGAGGTGGTGGTCTGGCCGGAGCGATCGGC 1140
L A G C E V K S G D E V V V L A G A I G

1141 CGGAACGGACCGTCCGCGACGCGCCCGCCCTGCCCGGCCAGCGGCCCGCCCGCC 1200
R N G P S A A A P P A P P G P A A P P A

1201 CCGTCGGTCTTCGGTGCCCGCCCTTCGAGAACGCGCTGGCCGAACCCCTCGTCCGGGCT 1260
P S V F G A A A F E N A L A E P L V R A

1261 GTGACGGGAGCGGCCCTCCAGGCCCTCGCGGAGGGGCCCGCCCGGCTGACGGCGGGGA 1320
V T G A A L Q A L A E G P P R L T A A G

1321 CCCGTCGTACGACGGCGGCTCCCTGTCTCGCGGGGCTGCACCGGGCTCCGGTGGCC 1380
P V V R R R R S P V V G G L H R A P V A

FIGURE 22

11381	GCCGCATGAGCATCGCGTCGAACGGCGCGCGCTCGGCCCGCCCGCGCCCTGCGCGTGA	1440
	A A *	V M
		OleG1
11441	TGATGACCACCTTCGGGCCAACACGCACTTCCAGCCGCTGGTTCCCTTGGCCTGGGCAC	1500
	M T T F A A N T H F Q P L V P L A W A L	
11501	TGCGGACAGCCGGGACGAGGTGGCGGTGGTGAGCCAGCCCTCGCTGAGCGACGTGGTGA	1560
	R T A G H E V R V V S Q P S L S D V V T	
11561	CGCAGGGGGGCTCACCTTCGGTCCCGGTGGGCACCGAGGCTCCGGTCGAGCAGTTCGCGG	1620
	Q A G L T S V P V G T E A P V E Q F A A	
11621	CGACCTGGGGCGACGATGCCCTACATCGGCGTCAACAGCATCGACTTCACCGGCAACGACC	1680
	T W G D D A Y I G V N S I D F T G N D P	
11681	CCGGCCTGTGGACGTGGCCGTACCTCCTGGGCATGGAGACCATGCTGGTGCCGGCCTTCT	1740
	G L W T W P Y L L G M E T M L V P A F Y	
11741	ACGAGTTGCTGAACAACGAGTCCCTTCGTGGACGGCGTAGTCGAGTTCGCCCGCTGACTGGC	1800
	E L L N N E S F V D G V V E F A R D W R	
11801	GGCCCGACCTGGTGATCTGGGAGCCGCTGACGTTCCGCCGGCGGTGGCGGCGCGGTCA	1860
	P D L V I W E P L T F A G A V A A R V T	
11861	CCGGCGGGCCCCACGCCCGCTGCCGTGGGGCAGGAGATCACCCCTGCGCGGGCGGCAGG	1920
	G A A H A R L P W G Q E I T L R G R Q A	
11921	CGTTCCCTCGCCGAGCGGTGCCCTGCNAACCGTTCGAGCACCGGAGGATCCCACGGCCGAGT	1980
	F L A E R A L Q P F E H R E D P T A E W	
11981	GGCTGGGGCGCATGCTCGACCCGGTACGGCTGCTCGTTCGACGAGGAGATGGTCAACCGGC	2040
	L G R M L D R Y G C S F D E E M V T G Q	

Bamili

FIGURE 22

FIGURE 22

3361 CGGCTGCCACCGGGACGACGAGTGGCGGTACGTGCCGTACACGGGGGGCC 3420
R L P T G T T V P M R Y V P Y N G R A

3421 GTGGTCCCGCATGGGTCCGGCAGCGTCCGGGGGGCCCGGATCTGCTGACGCTCGGT 3480
V V P A W V R Q R A R R P R I C L T L G

3481 GTGTCCGGCCCGGACAGACCTGGGGCAGCGCGTGTGGCGGAGGTGTGGCCGCGCTG 3540
V S A R Q T L G D G V S L A E V L A A L

3541 GCGACGTGGACCGGGAGATCGTGGCCACCGCTGGACCGCTCCACGCGCAAGCTCCTGGGG 3600
G D V D A E I V A T L D A S Q R K L L G

3601 CCGGTGCCGGACAAACGTCCGGCTGGTGGACTCTGTGCCCCCTGCACGCGCCCTGATGCCGACC 3660
P V P D N V R L V D F V P L H A L M P T

KpnI

3661 TGTTCGGCGCATCGTGCACCCACGCGGGCGCGGTACCTGGCTGACGGCCGCGCTCCACGGC 3720
C S A I V H H G G A G T W L T A A V H G

3721 GTCCCGCAGATCGTCTCGGTGACCTCTGGGACAACTGTGCGCGCCCGGCGAGACACAG 3780
V P Q I V L G D L W D N L L R A R Q T Q

3781 GCCGCGGGCGCGGCTGTTCATCCATCCGTCCGAGGTACCGCGGGCGGGCTCGGTGAG 3840
A A G A G L F I H P S E V T A A G L G E

3841 GCGTGGCGCGGTGTGACGAGACCTTCCATCCGGGGCGCGGACAGCGCGTCCGGGGAC 3900
G V R R V L T D P S I R A A A Q R V R D

3901 GAGATGAATGCAGAGCCGACGCGGGCGAGGTGCTACGGTGTGGAGCGGCTCGCCCGG 3960
E M N A E P T P G E V V T V L E R L A A

FIGURE 22

3961 AGCGCGGACGCGGACGAGGAGGGGGAACCATGCGGGGTGACACGGAGCCGACCAACCGG 4020
S G G R G R G G G N H A G *
M R A D T E P T T G
oleM

BgIII

4021 GTACGAGGACGAGTTCCGCCGAGATCTACGAGCCCGTGTACCGGGCGCGGCAAGGACTA 4080
Y E D E F A E I Y D A V Y R G R G K D Y

4081 CGCCGGCGAGGCGAAGGACGTGGCGGACCTCGTGGCGACCGGGTGCCGGACGCGTCCTC 4140
A G E A K D V A D L V R D R V P D A S S

4141 CCTCCTGGACGTGGCCCTGCGGCACGCGGGCGGCACCTGCGGCACTTCGCCACCGCTCTTCCA 4200
L L D V A C G T G A H L R H F A T L F D

4201 CGACGCGCGGGTCTCGAACTGTCCGCGAGCATGTGGACATCGCCCGCTCCCGCATGCC 4260
D A R G L E L S A S M L D I A R S R M P

4261 GGGCGTCCGCTGCACCAAGGGGACATGCGATCCTTCGACCTGGGGCCACGCGTCTCCGC 4320
G V P L H Q G D M R S F D L G P R V S A

4321 GGTACCTGCATGTTCAGCTCCGTGCGCCACCTGGCCACCAACCGCGCAACTCGACGCGAC 4380
V T C M F S S V G H L A T T A E L D A T

4381 GCTGCGGTGCTTCGCCCGGCACACCCGCGCGGTGGCGTGCATCGAACCGTGGTG 4440
L R C F A R H T R P G G V A V I E P W W

4441 GTTCCCGGAGACCTTCACCGACGGCTACGTGGCGGTGACATCGTACGCGTCGACGGCCG 4500
F P E T F T D G Y V A G D I V R V D G R

4501 GACCATCTCCCGGTGTCCCACTCGGTACGGGACGCGGGCGCCACCCCGCATGGAGATCCA 4560
T I S R V S H S V R D G G A T R M E I H

FIGURE 22

4561 CTACGTGATCGCCGACGCGGAGCACCGGTCGCCGGCACCTGGTCGAGCACCCACCGCATCAC 4620
Y V I A D A E H G P R H L V E H H R I T
4621 GCTGTTCGCCGGCATGCGTACACGCGCGCGGTACGAGAAGCGGGGCTACACCGTCGAGTA 4680
L F P R H A Y T A A Y E K A G Y T V E Y
4681 CCTCGACGGGGCCCTCGGGCGCGGGGCTGTTCTGTCGGCACCCCGGACGTGAACCCGCC 4740
L D G G P S G R G L F V G T R T *
4741 GCGCACCGCCCGATCACCCCTGCTCAACGCCCGTTCAACGCGATCACCGGACCCACGCGAAGG 4800
4801 ACCTTTCACATGTCGTACGACGACACCGCGGTGCTGGAAGCGATACCTGCGTGCGCCGGA 4860
M S Y D D H A V L E A I L R C A G
oley
4861 GGTGACGAGCGCTTCCTGCTGAACACCGTCGAGGAATGGGAGCGCGGAGATCACCGCG 4920
G D E R F L L N T V E E W G A A E I T A
4921 GCGCTCGTGACGAGTTGCTGTCCGCTGCGAGATCCCGCAGGTGGCGGTTGAGGCGTTC 4980
A L V D E L L F R C E I P Q V G G E A F
PstI
4981 ATCGGCCTGGACGTCCTGACGCGCGCCGACCGGATCAGCCATGTGCTGCAGGTGACGGAC 5040
I G L D V L H G A D R I S H V L Q V T D
5041 GGCAAGCCGGTCACGTCGGCGGAACCGCGCGCCAGGAACCTGGCGGCCGTACCTGGAGT 5100
G K P V T S A E P A G Q E L G G R T W S
5101 TCACGCTCAGCACCCCTCCTGCGGGAGCTGTTGCGCCCGCGCTCGGCCGCGCACCGCGGG 5160
S R S A T L L R E L F G P P S G R T A G
5161 GGCTTCGGCGTCTCCTTCTGCCCCGACCTGCGCGCGCCCGCGGACCATGAGGGCGCGGCC 5220
G F G V S F L P D L R G P R T M E G A A

FIGURE 22

5221 CTGGCCGCGCGCCACCAACGTGGTGTGCTGCACGGACGACCAACGAGACGCCCCCACTG 5280
L A A R A T N V V L H A T T N E T P P L

5281 GACCGGCTGGCCCTACGAGTCCGACAAAGTGGGGCGCGTCCACTGTTACCCGGC 5340
D R L A L R Y E S D K W G G V H W F T G

BamHI

5341 CACTACGACGGCACCTGCGGGCGCGTGGCGACGAGCGGTGCGGATCCTGGAGATCGGC 5400
H Y D R H L R A V R D Q A V R I L E I G

5401 ATCGCGGGCTACGACGACCTGCTGCCGAGCGGGCCCTCACTGAAGATGTGGAAGCGCTAC 5460
I G G Y D D L L P S G A S L K M W K R Y

5461 TTCCCGCGGGCCCTGGTCTTCGGCGTGGACATCTTCGACAGTGGCGGTGCGACCGCCGC 5520
F P R G L V F G V D I F D S R R A T S R

5521 GTGTCAAGACGCTCCGCGGGCCCGGACGAGGACGCCGGAGTTCATGGCGCGCGTCCGCCGAG 5580
V S R R S A A R Q D D P E F M R R V A E

5581 GAGCACGGGCGGTTCGACGTCATCATCGACGCGGCGGACCCACATCAACGCACACATGCCG 5640
E H G P F D V I I D D G S H I N A H M R

5641 ACGTCGTTCTCGGTGATGTTCCCCCACCCTGGCACAACGGCGGCTTCTACGTCATCGAGGAC 5700
T S F S V M F P H L R N G G F Y V I E D

5701 ACCTTCACCTCCTACTGGCCCGGGTACGGAGGGCCCATCCGGAGCCCCGGTCCCCGTCCGGA 5760
T F T S Y W P G Y G G P S G A R C P S G

5761 ACAACCGGCTGGAGATGGTCAAGGACTGATCGACTCGGTGCACCTACGAGGAGCGGCCG 5820
T T A L E M V K G L I D S V H Y E E R P

FIGURE 22

5821 GACGGCGGGCCACGGGCGGACTACATCGCCAGGAACCTCGTCGGGGCTGCACGGCCTACCAA 5880
D G A A T A D Y I A R N L V G L H A Y Q

5881 ACGACCTCGTCTTCCTCGAGAGGGGCGGATCAACAAGAGGGGGCGCATCCCCCACAACCGTG 5940
T T S S S R R A I N K E G G I P H T V

5941 CCCCCGGAGCCGTTCTGTGAACGACAACCTAGCCACGGCCGCAACCAGAGCCCGGAAACCGCA 6000
P R E P F W N D N *

6001 CCACTGTCCGGCCACCTCGGAACCCACCTCCAGCAAGGACACACCCGCTGTGACCGGATAC 6060

6061 GCACACGGACCGACACCGGCGCGGTACC 6093

KpnI

2476/PCT/05

Please type a plus sign (+) inside this box → ☐

PTO/SB/01 (8-96)

Approved for use through 9/30/98. OMB 0651-0032

Patent and Trademark Office: U.S. DEPARTMENT OF COMMERCE

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

DECLARATION FOR UTILITY OR DESIGN PATENT APPLICATION

☒ Declaration OR
Submitted
with Initial Filing

☐ Declaration
Submitted after
Initial Filing

Attorney Docket Number	146.1335
First Named Inventor	JESUS CORTES
COMPLETE IF KNOWN	
Application Number	PCT/FR98/01593
Filing Date	July 21, 1998
Group Art Unit	
Examiner Name	

As a below named inventor, I hereby declare that:

My residence, post office address, and citizenship are as stated below next to my name.

I believe I am the original, first and sole inventor (if only one name is listed below) or an original, first and joint inventor (if plural names are listed below) of the subject matter which is claimed and for which a patent is sought on the invention entitled:

BIOSYNTHESIS GENES AND TRANSFER OF 6-DESOXY-HEXOSES IN
SACCAHAROPOLYSPORA ERYTHRAEA AND IN STREPTOMYCES ANTIOBIOTICUS
AND THEIR USE

(Title of the Invention)

the specification of which

☐ is attached hereto
OR

☒ was filed on (MM/DD/YYYY) July 21, 1998

as United States Application Number or PCT International

Application Number PCT/FR98/01593 and was amended on (MM/DD/YYYY) (if applicable).

I hereby state that I have reviewed and understand the contents of the above identified specification, including the claims, as amended by any amendment specifically referred to above.

I acknowledge the duty to disclose information which is material to patentability as defined in Title 37 Code of Federal Regulations, §1.56.

I hereby claim foreign priority benefits under Title 35, United States Code §119 (a)-(d) or §365(b) of any foreign application(s) for patent or inventor's certificate, or §365 (a) of any PCT international application which designated at least one country other than the United States of America, listed below and have also identified below, by checking the box, any foreign application for patent or inventor's certificate, or of any PCT international application having a filing date before that of the application on which priority is claimed.

Prior Foreign Application Number(s)	Country	Foreign Filing Date (MM/DD/YYYY)	Priority Not Claimed	Certified Copy Attached?	
				YES	NO
97/09458	FRANCE	Jul. 25, 97	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
98/07411	FRANCE	Jun. 12, 98	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
PCT/FR98/01593	FRANCE	Jul. 21, 98	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
			<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
			<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
			<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

☐ Additional foreign application numbers are listed on a supplemental priority sheet attached hereto:

I hereby claim the benefit under Title 35, United States Code §119(e) of any United States provisional application(s) listed below.

Application Number(s)	Filing Date (MM/DD/YYYY)	<input type="checkbox"/> Additional provisional application numbers are listed on a supplemental priority sheet attached hereto.

[Page 1 of 5]

Burden Hour Statement: This form is estimated to take 0.4 hours to complete. Time will vary depending upon the needs of the individual case. Any comments on the amount of time you are required to complete this form should be sent to the Chief Information Officer, Patent and Trademark Office, Washington, DC 20231. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. SEND TO: Commissioner of Patents and Trademarks, Washington, DC 20231.

(January 1997)

Please type a plus sign (+) inside the box →


 Approved for use through 9/30/98 OMB 0351-0032
 Patent and Trademark Office, U.S. DEPARTMENT OF COMMERCE
 Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

DECLARATION

I hereby claim the benefit under Title 35, United States Code §120 of any United States application(s), or §365(c) of any PCT international application designating the United States of America, listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in the prior United States or PCT International application in the manner provided by the first paragraph of Title 35, United States Code §112, I acknowledge the duty to disclose information which is material to patentability as defined in Title 37, Code of Federal Regulations §1.56 which became available between the filing date of the prior application and the national or PCT international filing date of this application.

U.S. Parent Application Number	PCT Parent Number	Parent Filing Date (MM/DD/YYYY)	Parent Patent Number (if applicable)

☐ Additional U.S. or PCT international application numbers are listed on a supplemental priority sheet attached hereto.

As a named inventor, I hereby appoint the following registered practitioner(s) to prosecute this application and to transact all business in the Patent and Trademark Office connected therewith:

Name	Registration Number	Name	Registration Number
Charles A. Muserlian	19,683	3	
Jordan B. Bierman	18,629		
Donald C. Lucas	31,275		
Bierman, Muserlian and Lucas	18,818		

☐ Additional registered practitioner(s) named on a supplemental sheet attached hereto.

Direct all correspondence to:

Name	Bierman, Muserlian and Lucas		
Address			
Address	600 Third Avenue		
City	New York	State	New York
Country	U.S.A.	Telephone	(212) 661-8000
		Fax	(212) 661-8002

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

Name of Sole or First Inventor:

☐ A petition has been filed for this unsigned inventor

Given Name	Jesus	Middle Initial		Family Name	CORTES	Suffix e.g. Jr.	
------------	-------	----------------	--	-------------	--------	-----------------	--

Inventor's Signature	<i>Jesus Cortes B</i>	Date	27/01/00
----------------------	-----------------------	------	----------

Residence: City	Cambridge	State		Country	United Kingdom	Citizenship	Mexican
-----------------	-----------	-------	--	---------	----------------	-------------	---------

Post Office Address	
---------------------	--

Post Office Address	26 Cambanks, Union Lane
---------------------	-------------------------

City	Cambridge	State		Zip	CB4 IPZ	Country	United Kingdom
------	-----------	-------	--	-----	---------	---------	----------------

☒ Additional inventors are being named on supplemental sheet(s) attached hereto

Please type a plus sign (+) inside this box → ☐

PTO/SB/01 (8-96)
Approved for use through 9/30/98. OMB 0651-0032
Patent and Trademark Office U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number

DECLARATION	ADDITIONAL INVENTOR(S) Supplemental Sheet
--------------------	--

Name of Additional Joint Inventor, if any:				<input type="checkbox"/> A petition has been filed for this unsigned inventor			
Given Name	Claude	Middle Initial		Family Name	FROMENTIN	Suffix e.g. Jr.	
Inventor's Signature					Date		
Residence: City	Paris	State		Country	France FRX	Citizenship	French
Post Office Address							
Post Office Address		16, Rue de Flandres					
City	Paris	State		Zip	F-75019	Country	France
Name of Additional Joint Inventor, if any:				<input type="checkbox"/> A petition has been filed for this unsigned inventor			
Given Name	Sabine	Middle Initial		Family Name	GAISSER	Suffix e.g. Jr.	
Inventor's Signature	<i>Sabine Gaiser</i>				Date	23.01.00	
Residence: City	Cambridge	State		Country	United Kingdom	Citizenship	German
Post Office Address							
Post Office Address		37 Gwydir Street					
City	Cambridge	State		Zip	CBI 2LG	Country	United Kingdom
Name of Additional Joint Inventor, if any:				<input type="checkbox"/> A petition has been filed for this unsigned inventor			
Given Name	Peter	Middle Initial		Family Name	LEADLAY	Suffix e.g. Jr.	
Inventor's Signature	<i>P. Leadlay</i>				Date	24 Jan 2000	
Residence: City	Cambridge	State		Country	United Kingdom	Citizenship	English
Post Office Address							
Post Office Address		17 Clarendon Road					
City	Cambridge	State		Zip	CB2 2BH	Country	United Kingdom
Name of Additional Joint Inventor, if any:				<input type="checkbox"/> A petition has been filed for this unsigned inventor			
Given Name	Carmen	Middle Initial		Family Name	MENDEZ	Suffix e.g. Jr.	
Inventor's Signature					Date		
Residence: City	Oviedo	State		Country	Spain ESX	Citizenship	Spanish
Post Office Address							
Post Office Address		C/Marcelino Fernandez 7, 2ºB					
City	Oviedo	State		Zip	ES-33010	Country	Spain
<input type="checkbox"/> Additional inventors are being named on supplemental sheet(s) attached hereto							

Please type a plus sign (+) inside this box → ☐

PTO/SB/01 (8-96)
Approved for use through 9/30/98. OMB 0651-0032
Patent and Trademark Office: U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

DECLARATION	ADDITIONAL INVENTOR(S) Supplemental Sheet
--------------------	--

Name of Additional Joint Inventor, if any:				<input type="checkbox"/> A petition has been filed for this unsigned inventor			
Given Name	Jean-Marc	Middle Initial		Family Name	MICHEL	Suffix e.g. Jr.	
Inventor's Signature					Date		
Residence: City	Compiègne	State		Country	France FRX	Citizenship	French
Post Office Address							
Post Office Address	22, Rue des Domeliers						
City	Compiègne	State		Zip	F-60200	Country	France
Name of Additional Joint Inventor, if any:				<input type="checkbox"/> A petition has been filed for this unsigned inventor			
Given Name	Marie-Cecile	Middle Initial		Family Name	RAYNAL	Suffix e.g. Jr.	
Inventor's Signature					Date		
Residence: City	Paris	State		Country	France FRX	Citizenship	French
Post Office Address							
Post Office Address	117, Avenue de Choisy						
City	Paris	State		Zip	F-75013	Country	France
Name of Additional Joint Inventor, if any:				<input type="checkbox"/> A petition has been filed for this unsigned inventor			
Given Name	Khadidja	Middle Initial		Family Name	SALAH-BEY	Suffix e.g. Jr.	
Inventor's Signature					Date		
Residence: City	Paris	State		Country	France FRX	Citizenship	Algerian
Post Office Address							
Post Office Address	100, Boulevard Massena						
City	Paris	State		Zip	F-75013	Country	France
Name of Additional Joint Inventor, if any:				<input type="checkbox"/> A petition has been filed for this unsigned inventor			
Given Name	Jose	Middle Initial		Family Name	SALAS	Suffix e.g. Jr.	
Inventor's Signature					Date		
Residence: City	Oviedo	State		Country	Spain ESX	Citizenship	Spanish
Post Office Address							
Post Office Address	C/Guillermo Estrada 2-Bajo Izda						
City	Oviedo	State		Zip	ES-33060	Country	Spain
<input type="checkbox"/> Additional inventors are being named on supplemental sheet(s) attached hereto							